

**ADVANCED PROTEOMIC CHARACTERIZATION OF THE 26S PROTEASOME IN
ARABIDOPSIS REVEALS INSIGHTS INTO COMPOSITION AND ASSEMBLY**

by

David C. Gemperline

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Genetics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2016

Date of final oral examination: TBD

The dissertation is approved by the following members of the Final Oral Committee:

Richard D. Vierstra, Professor, Genetics

Richard Amasino, Professor, Biochemistry

Josh Coon, Professor, Chemistry

Donna Fernandez, Professor, Botany

Patrick Masson, Professor, Genetics

© Copyright by David C. Gemperline 2016

All Rights Reserved

To Erin, my wife.

ACKNOWLEDGMENTS

Science doesn't purvey absolute truth. Science is a mechanism, a way of trying to improve your knowledge of nature. It's a system for testing your thoughts against the universe and seeing whether they match. This works not just for the ordinary aspects of science, but for all of life.

— ISAAC ASIMOV (1988)

Acknowledgements go here.

TABLE OF CONTENTS

Table of Contents	iii
List of Tables	iv
List of Figures	v
List of Abbreviations and Acronyms	vi
Abstract	vii
Abstract	ix
Chapter 1: Proteomic Characterization of Affinity-Purified Arabidopsis Pro-	
teosome Particles Identifies Sub-Complex Specific Interactors	1
1.1 Statement of Attribution	1
1.2 Abstract	2
1.3 Introduction	4
1.4 Experimental Procedures	10
1.4.1 Transgenic Plants and Growth Conditions	10
1.4.2 Sequence and Phylogenetic Analyses	13

1.4.3	Proteasome Affinity Purifications using PAG1-FLAG and FLAG-RPT4a/b	14
1.4.4	SDS-PAGE and Native-PAGE Analysis and Silver Staining	15
1.4.5	MS/MS Sample Preparation	17
1.4.6	MS/MS Analyses	19
1.4.7	MS/MS Data Processing	20
1.4.8	Yeast-two hybrid, bimolecular fluorescent complementation, and HA-PAP1 immunopurification	23
1.5	Results	25
1.5.1	<i>Arabidopsis</i> Proteasomes Can Be Effectively Affinity-Purified Using Epitope-Tagged Versions of Either Isoform of the Reg- ulatory Particle Subunit RPT4	25
1.5.2	MS/MS Analyses of RPT4a/b Affinity-Purified Proteins Iden- tifies a Similar Complement of Proteins as PAG1-FLAG Affin- ity Purifications	32
1.5.3	Statistical Analysis of Affinity Enrichments for FLAG-RPT4a Show Little to No Difference When Compared to FLAG-RPT4b	37

1.5.4 Statistical Analyses of LFQ-MS/MS Data Show Affinity Preparations Targeting Either the CP or RP Specifically Enrich for Their Respective Subcomplexes	38
1.5.5 Statistical Analysis of Affinity Purifications Enriched for the RP or CP Identifies Proteasome-Associated Proteins and Putative Assembly Chaperones Specific to Each Subcomplex	43
1.5.6 A Novel Plant Proteasome-Associated Protein, PAP1, Interacts with a Putative CP Assembly Chaperone	46
1.5.7 Inhibition of the Proteasome Results in Formation of Distinct Proteasome Sub-Species Containing Putative CP Assembly Chaperones	50
1.6 Discussion	56
1.7 Literature Cited	71
Colophon	111

LIST OF TABLES

3.S1 Oligonucleotide Primer Sequences Used in this Study	104
3.S2 Contaminants Removed from MaxQuants ProteinGroups.txt file	105
3.S3 InterProScan analysis of <i>Arabidopsis</i> protein sequences for PBAC1-4 . .	106
3.S4 Iterative PSI-BLAST Analysis using Human PAC1-4 to Identify PBAC1-4	107
3.S5 PAG1 Associated	110

LIST OF FIGURES

1.1 <i>rpt4a-1</i> and <i>rpt4b-2</i> mutants rescued with flag tagged variants	26
1.2 Protein composition of affinity-purified proteasomes from FLAG-RPT4a-, FLAG-RPT4b-, and PAG1-FLAG-expressing plants and from mock wild- type controls in the presence (+) or absence (-) of 20mM ATP	31
1.3 Label-free quantitative MS/MS analysis of affinity-purified proteins in the presence (+) or absence (-) of ATP	34
1.3 Label-free quantitative MS/MS analysis of affinity-purified proteins in the presence (+) or absence (-) of ATP	35
1.3 Label-free quantitative MS/MS analysis of affinity-purified proteins in the presence (+) or absence (-) of ATP	36
1.4 Volcano Plots comparing proteasome affinity purifications under both plus and minus ATP conditions relative to WT controls, allow identifica- tion of statistically significant interacting partners for the proteasome .	40
1.4 Volcano Plots comparing proteasome affinity purifications under both plus and minus ATP conditions relative to WT controls, allow identifica- tion of statistically significant interacting partners for the proteasome .	41

1.5 PAP1 interacts with the proteasome and with putative assembly chaperone PBAC1	49
1.6 Protein composition of PAG1-FLAG affinity-purified proteasomes from tissues treated with and without MG132	52
1.6 Protein composition of PAG1-FLAG affinity-purified proteasomes from tissues treated with and without MG132	53
1.7 Putative Model of 20S Proteasome Assembly in <i>Arabidopsis</i>	69
3.S1 Sequence Alignments of RPT4a and RPT4b	80
3.S2 Volcano Plots Comparing Protein Profiles of Proteasome Affinity Purifications Based on RPT4a or RPT4b subunits in the presence (+) or absence (-) of ATP	81
3.S3 Kernel Density Estimates (KDE) of the CP and RP distributions from label-free quantitative mass spectrometry show enrichment for the CP when purified via PAG1-FLAG, and for the RP when purified via FLAG-RPT4a/b	82
3.S4 Multiple Sequence Alignment for UMP1	83
3.S4 Multiple Sequence Alignment for UMP1	85
3.S5 Multiple Sequence Alignment for PBAC1	86
3.S5 Multiple Sequence Alignment for PBAC1	88

3.S6 Multiple Sequence Alignment for PBAC2	89
3.S6 Multiple Sequence Alignment for PBAC2	91
3.S7 Multiple Sequence Alignment for PBAC3	92
3.S7 Multiple Sequence Alignment for PBAC3	94
3.S8 Multiple Sequence Alignment for PBAC4	95
3.S8 Multiple Sequence Alignment for PBAC4	97
3.S9 Multiple Sequence Alignment for PAP1	99
3.S10 Full BiFC Panel and Controls	100
3.S11 Phylogenetic Analysis of PAP1 and PBAC1-4	101
3.S11 Phylogenetic Analysis of PAP1 and PBAC1-4.	103

LIST OF ABBREVIATIONS AND ACRONYMS

API	Application programming interface
BCA	Bicinchoninic acid protein assay
BLAST	Basic local alignment search tool
C#	C sharp, a programming language
Da	Dalton, the atomic mass unit
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ESI	Electrospray ionization
E-value	Expectation value
FASTA	A format for storing protein sequences
FDR	False discovery rate
GUI	Graphical user interface
HCD	Higher-energy collisional dissociation
HPLC	High-performance liquid chromatography
LC	Liquid chromatography
<i>m</i>	Mass
min	Minute
MS	Mass spectrometry

MS ¹	Survey mass analysis
MS/MS	Tandem mass spectrometry
NCE	Normalized collision energy
nLC	Nanoflow liquid chromatography
ppm	Part per million
PSM	Peptide-spectrum match
PTM	Post-translational modification
s	Second
SILAC	Stable isotope labeling by amino acids in cell culture
S/N	Signal-to-noise ratio
TMT	Tandem mass tag

**ADVANCED PROTEOMIC CHARACTERIZATION OF THE 26S
PROTEASOME IN ARABIDOPSIS REVEALS INSIGHTS INTO
COMPOSITION AND ASSEMBLY**

David C. Gemperline

Under the supervision of Professor Richard D. Vierstra

At the University of Wisconsin-Madison

Abstract

The 26S proteasome is the central proteolytic effector in the ubiquitin system that is responsible for degrading numerous regulators following their selective ubiquitylation. While much is known about the construction of the yeast and mammal particles, little is known about the pathways used to assemble the plant particle. One challenge is that the known yeast chaperones appear sufficiently diverged to preclude high-confidence identification of their plant counterparts by genomic searches. Here, we used in-depth mass spectrometric analysis of Arabidopsis 26S proteasomes, which were affinity purified from seedlings under conditions that promote the accumulation of assembly intermediates, to identify a large collection of interacting proteins that associate with either the core protease (CP) or regulatory particle (RP). Sequence comparisons, Y2H and BiFC studies revealed

that some are likely assembly chaperones, with several CP factors harboring the signature C-terminal HbYX motif that allows their association with the α -subunit ring. Several of the RP-specific factors appear to be orthologs of the chaperones Nas2, Nas6, Hsm3 and Ecm29. Whereas yeast assembles only a single particle type, mammals can assemble alternate proteasomes by replacing individual subunits with distinct isoforms (e.g., immunoproteasomes). In plants, most 26S proteasome subunits are encoded by paralogous genes with sufficient divergence to suggest that plants also accumulate a collection of particles. However, proteomic analysis of proteasomes selectively enriched using paralog-specific tags strongly imply that although plants possess this genetic diversity, the incorporation of these paralogs appears random, and is mainly influenced by the differential expression of the corresponding genes. Taken together, these proteomic studies provide the first insights into plant proteasome assembly and diversity, and identify factors that build the CP and RP subcomplexes and finally the 26S holo-particle.

Richard D. Vierstra

ABSTRACT

Abstract

The 26S proteasome is the central proteolytic effector in the ubiquitin system that is responsible for degrading numerous regulators following their selective ubiquitylation. While much is known about the construction of the yeast and mammal particles, little is known about the pathways used to assemble the plant particle. One challenge is that the known yeast chaperones appear sufficiently diverged to preclude high-confidence identification of their plant counterparts by genomic searches. Here, we used in-depth mass spectrometric analysis of *Arabidopsis* 26S proteasomes, which were affinity purified from seedlings under conditions that promote the accumulation of assembly intermediates, to identify a large collection of interacting proteins that associate with either the core protease (CP) or regulatory particle (RP). Sequence comparisons, Y2H and BiFC studies revealed that some are likely assembly chaperones, with several CP factors harboring the signature C-terminal HbYX motif that allows their association with the α -subunit ring. Several of the RP-specific factors appear to be orthologs of the chaperones Nas2, Nas6, Hsm3 and Ecm29. Whereas yeast assembles only a single particle type, mammals can assemble alternate proteasomes by replacing individual subunits with distinct isoforms (e.g., immunoproteasomes). In plants, most 26S proteasome

subunits are encoded by paralogous genes with sufficient divergence to suggest that plants also accumulate a collection of particles. However, proteomic analysis of proteasomes selectively enriched using paralog-specific tags strongly imply that although plants possess this genetic diversity, the incorporation of these paralogs appears random, and is mainly influenced by the differential expression of the corresponding genes. Taken together, these proteomic studies provide the first insights into plant proteasome assembly and diversity, and identify factors that build the CP and RP subcomplexes and finally the 26S holo-particle.

Chapter 1

PROTEOMIC CHARACTERIZATION OF AFFINITY-PURIFIED ARABIDOPSIS PROTEASOME PARTICLES IDENTIFIES SUB-COMPLEX SPECIFIC INTERACTORS

1.1. Statement of Attribution

Initial mutant characterization of *rpt4* mutants was performed by Drs. Kwang-Hee Lee and Richard S. Marshall, postdoctoral fellows in the Vierstra Lab. Additionally, Dr. Lee generated the 2XFLAG-RPT4b transgenic line in the *rpt4b-2* background. I created the 2XFLAG-RPT4a *rpt4a-1* transgenic lines. I developed the affinity purification via the RP with minor modifications from (Book *et al.*, 2010) by changing the primary buffer to HEPES. All mass-spectrometry sample preparation and data analysis in this chapter was completed by myself. The actual running of the mass spectrometer was either performed by Dr. Mark Scalf in Dr. Lloyd M. Smiths lab (Department of Chemistry, University of Wisconsin Madison), or by Dr. Fionn McLoughlin in the Vierstra lab at Washington University. I cloned the CP interactors PAP1, PBAC1, PBAC2, PBAC3, and PBAC4, and performed yeast-two-hybrid (Y2H) analysis on these. While the streaking and yeast transformations were ac-

complished by myself, Dr. Marshall performed the final spotting of the yeast-two hybrid. All BiFC data was generated by myself. The HA-PAP1 transgenic line was created by Dr. Adam Book. I developed the affinity purification for HA-PAP1 and performed the immunoblot analyses. Native PAGE was performed by Dr. Marshall.

We intend to submit this chapter to *J. Biol. Chem.* for publication as a research article in the near future. A suggested list of authors is as follows: David Gemperline, Richard S. Marshall, Kwang-Hee Lee, Fionn McLoughlin, Mark Scalf, Lloyd M. Smith, and Richard D. Vierstra.

1.2. Abstract

The 26S proteasome is an ATP-dependent protease complex responsible for the selective removal of numerous eukaryotic proteins following their modification with polyubiquitin chains. It consists of two sub-complexes, a self-compartmentalized core protease (CP) that degrades polypeptides within a central chamber, and a regulatory particle (RP) that caps one or both ends of the CP and is responsible for recognizing and unfolding ubiquitylated substrates, and transporting them into the CP lumen. Previous mass spectrometric (MS) analyses of *Arabidopsis* 26S proteasomes affinity-purified via the CP subunit PAG1 ($\alpha 7$), detected a complex

assortment of particles that are assembled using paired isoforms for many of the 33 core subunits, as well as a small assortment of associated factors that might help in assembly, regulation, substrate specificity, and ubiquitin recycling. Here, we extended these studies by affinity purifying *Arabidopsis* 26S proteasomes with either tagged isoform of the RP subunit RPT4 (named RPT4a and RPT4b). Label-free quantitative (LFQ) MS analyses of the preparations detected both isoforms for many of the other CP and RP subunits, suggesting that plants incorporate subunit isoforms randomly into the holo-protease. We also identified an array of interacting factors, including likely orthologs of the yeast assembly chaperones Ump1 and Pba1-4 for the CP, and Nas2, Nas6, and Hsm3 for the RP. We additionally identified a novel plant-specific protein bound to the CP that interacts with the putative *Arabidopsis* Pba1 ortholog, and possibly contributes to CP assembly. By coupling native PAGE and MS analysis with proteasome preparations isolated from proteasome inhibitor-treated seedlings, we identified several assembly intermediates also harboring assembly chaperones and the proteasome regulators PA200 and ECM29. Taken together these data suggest that *Arabidopsis*, like yeast, dynamically employ a suite of assembly factors to construct the mature 26S particle.

1.3. Introduction

The selective removal of proteins that are non-functional or have key regulatory roles is essential to normal plant growth and development. A major pathway for this selective removal is the ubiquitin 26S proteasome system, in which target proteins are covalently modified with the 76-amino-acid polypeptide ubiquitin and subsequently degraded by the central enzymatic effector in this process, the 26S proteasome (reviewed in (Finley, 2009; Livneh *et al.*, 2016; Vierstra, 2009)). The 26S proteasome consists of two major sub-complexes, the core protease (CP), which is responsible for degrading the substrate, and the regulatory particle (RP), which caps the CP at one or both ends and is responsible for recognizing and unfolding ubiquitylated substrates. Compositionally, the CP is made up of 28 subunits configured into four stacked hetero-heptameric rings made up of α and β -subunits in an α 1-7/ β 1-7/ β 1-7/ α 1-7 arrangement. The N-termini of the β 1, β 2, and β 5 subunits are enclosed in a central chamber by the surrounding α -subunits, and together these β -subunits provide peptidylglutamyl-peptide-hydrolyzing activity (β 1), trypsin-like (β 2), and chymotrypsin-like (β 5) activities. The α -subunits possess N-terminal extensions that gate entry to the chamber, forming an axial pore that is only opened when substrates are actively processed by an associated RP, or when activated by alternative capping particles such as PA200 (Dange *et al.*, 2011;

Sadre-Bazzaz *et al.*, 2010).

The RP consists of two major sub-classes of proteins including six related AAA-ATPase (RPT) subunits (RPT1-6) and at least 15 non-AAA-ATPase (RPN) subunits (RPN1-3, 5-13, and SEM1) (Finley, 2009; Paraskevopoulos *et al.*, 2014; Russell *et al.*, 2013). The RPT subunits form a six-membered RPT ring responsible for unfolding the polypeptide substrate in an ATP dependent manner, and channeling the substrate through the axial pore into the central CP chamber. The RPN subunits have several differing functions including binding ubiquitylated substrates (RPN1, RPN10, RPN13, and SEM1 (Elsasser *et al.*, 2004; Paraskevopoulos *et al.*, 2014; Schreiner *et al.*, 2008; Shi *et al.*, 2016)), binding ubiquitin shuttle proteins such as DSK2 and RAD23 (Elsasser *et al.*, 2002; Farmer *et al.*, 2010; Fatimababy *et al.*, 2010; Lin *et al.*, 2011), and releasing ubiquitin bound to substrates through the deubiquitylating activity of RPN11 (Verma *et al.*, 2002; Yao and Cohen, 2002). RPN5-7 and RPN9 form a glove-like structure around the RPT ring providing structural support for adjacent subunits (Lander *et al.*, 2012; Lasker *et al.*, 2012; Unverdorben *et al.*, 2014), with RPN6 acting as a molecular clamp that holds the RP and CP together (Pathare *et al.*, 2012).

Over the last several years, analyses of the plant 26S proteasome have also revealed insights into its composition. Searches of the *Arabidopsis* genome showed

that plants typically encode two paralogs for most proteasome subunits (Fu *et al.*, 1998). Almost all of these paralogs were identified in purified preparations of the *Arabidopsis* proteasome that used either chromatographic strategies (Yang *et al.*, 2004), or an affinity purification method that exploited a FLAG tagged CP subunit PAG1 (Book *et al.*, 2010). Like in *Arabidopsis*, duplication of RPT genes have also occurred in monocots, with rice having duplicated genes for RPT1, 2, 4 and 5, suggesting that there may be some evolutionary benefit to having multiple copies of the RPT family (Shibahara *et al.*, 2004). Taken together these data suggest that plants have the capacity to generate a highly diverse set of proteasome isotypes by preferentially combining specific subsets of isoforms with others into the holo particle. Such a precedent for proteasome isotypes exist in mammals, which assemble the immunoproteasome and thymoproteasome isotypes by selectively incorporating alternative isoforms of the catalytic β -subunits (Murata *et al.*, 2007; Nandi *et al.*, 1996). Similarly, the mammalian testes proteasome isotype incorporates an alternative $\alpha 4$ isoform (Belote *et al.*, 1998; Uechi *et al.*, 2014). In plants, genetic evidence suggests that some subunit paralogs may have distinct functions. For instance, RPN1a and RPN1b play differing roles in embryogenesis (Brukhin *et al.*, 2005). Furthermore, overexpression of RPN5a induces an early senescent phenotype while overexpression of RPN5b does not (Book *et al.*, 2009). Despite these observations, it

remains unclear if plants deliberately assemble their proteasome subunit isoforms into compositionally and functionally distinct proteasome isotypes.

Outside of the core subunits, other proteins are known to associate with the proteasome transiently or sub-stoichiometrically. These proteasome-associated proteins, or PAPs, perform a variety of functions including, substrate processing, ubiquitin recycling, and proteasome assembly. DSK2 and RAD23 contribute to the shuttling of ubiquitylated substrates to the proteasome (Farmer *et al.*, 2010; Fatimababy *et al.*, 2010; Lin *et al.*, 2011), RPN13 helps bind ubiquitylated targets (Schreiner *et al.*, 2008), UCH37 and other de-ubiquitylating enzymes (DUBs) (Van-derLinden *et al.*, 2015) help process substrates bound to the RP, and a suite of assembly chaperones assist in the construction of both the CP and RP. While the route(s) for plant proteasome assembly are largely unknown, a possible model has been established in yeast and mammals (see Figure 1.7, and Figure 1.10). In both systems, the Pba3 chaperone (PAC3 in mammals) interacts with Pba4 (PAC4), forming a Pba(PAC)3/4 heterodimer aiding in early stages of CP assembly by binding the α 5 subunit, and likely adjacent α -subunits (Kunjappu and Hochstrasser, 2014; Yashiroda *et al.*, 2008). Similarly, Pba1 (PAC1) interacts with Pba2 (PAC2) to form a Pba(PAC)1/2 heterodimer that aide in assembly of the α -ring, and subsequently together with Ump1, helps to form 15S half-barrels consisting of α 1-7/ β 1-6 notably

lacking the $\beta 7$ subunit (Kunjappu and Hochstrasser, 2014; Marques *et al.*, 2007). While the Pba(PAC)1/2 heterodimer act throughout assembly and can even bind mature CP (Stadtmueller *et al.*, 2012), the Pba(PAC)3/4 heterodimer leaves upon incorporation of the $\beta 3$ subunit prior to the formation of 15S half-barrels, thus only acting at early stages of assembly (Hirano *et al.*, 2008). Finally, CP maturation occurs concomitant with autolytic cleavage of β -subunit propeptide-appendages exposing proteolytic N-terminal threonines and subsequent degradation of Ump1, the first substrate of the CP (Ramos *et al.*, 1998). In addition to these dedicated CP assembly chaperones, in humans and yeast a different suite of other chaperones also guides assembly of the RP, including Nas2, Nas6, Hsm3, and Rpn14 (see Figure 1.10). These RP assembly chaperones typically bind to the C-terminal domains of specific RPT subunits forming modules (Nas2 with Rpt4 and 5, Nas6 with Rpt3, Hsm3 with Rpt1,2 and Rpn1, and Rpn14 with Rpt6) that, importantly, prevent the assembling RPT ring from forming a pre-mature CP-RP interface (Park *et al.*, 2010). In plants, MS/MS analyses of *Arabidopsis* proteasomes identified only one putative CP assembly chaperone, PBAC2 (the nomenclature used here is different from yeast (Pba) and mammals (PAC) as both are standard nomenclature for the $\beta 1$ and $\alpha 1$ subunits in *Arabidopsis*) (Book *et al.*, 2010), which is an ortholog to PAC2 based on PSI-BLAST sequence analyses (Kusmierczyk *et al.*, 2011; Le Tallec *et al.*, 2007).

Together these data suggest that other plant CP assembly chaperones remain to be identified (Book *et al.*, 2010). Additionally, these proteomic analyses failed to identify obvious plant RP assembly orthologs, such as NAS2, NAS6 and HSM3, despite the identification of these orthologs in the *Arabidopsis* genome (Book *et al.*, 2010). Taken together, the current data suggest that our understanding of plant PAPs in general, and CP and RP assembly chaperones in particular, remains incomplete.

While tools exist in other organisms to affinity-purify the 26S proteasome via the RP, including Pro-A tagged variants of RPN11, and RPT1 in yeast (Leggett *et al.*, 2005, 2002), and RAD23-based strategies that bind RPN1 in mammals (Besche *et al.*, 2009), such systems were not available in plants until this work. We also reasoned that our existing CP-based affinity purification might miss some RP-specific PAPSs. To overcome this challenge and to explore the possibility of plants forming proteasome isotypes, we sought to develop an RP affinity-purification strategy using two FLAG epitope-tagged RP subunit isoforms, RPT4a and RPT4b. RPT4 was specifically chosen for affinity tagging as prior structural analyses indicated that its N-terminus was likely solvent exposed (PDB 4CR2) (Beck *et al.*, 2012). We also exploited the ATP dependence of the interaction between the CP and RP to enrich for both of these sub-complexes specifically (Book *et al.*, 2010; Liu *et al.*, 2006). These approaches together allowed us to discover novel CP and RP interacting partners, including

several putative assembly chaperones. Follow-up interaction studies imply a role for several of these associated proteins in CP assembly. Finally, proteomic analyses of native-gel separated complexes from tissues treated with the potent and specific proteasome inhibitor MG132 identified novel sub-complexes that contain distinct subgroups of these putative assembly chaperones, which presumably represent assembly intermediates. From these results, we propose a model that assigns several of these novel associated proteins with distinct phases of proteasome assembly.

1.4. Experimental Procedures

1.4.1. Transgenic Plants and Growth Conditions

The *rpt4a-1*, *rpt4a-3*, *rpt4a-4*, and *rpt4b-1*, *rpt4b-2*, and *rpt4b-3* T(transfer)-DNA insertion mutants in the *Arabidopsis thaliana* Columbia-0 ecotype (Col-0) (SALK_052372, SALK_128087C, SALK_135246, SALK_108556, SALK_108557, and SALK_101982C, respectively) were obtained from the *Arabidopsis* Biological Resource Center (The Ohio State University, Columbus, OH). All mutants were backcrossed three times to the Col-0 parent and then made homozygous by self-fertilization. The *PAG1::PAG1-FLAG pag1-1* germplasm was as previously described (Book *et al.*, 2010). The *rpt4a* and *rpt4b* alleles were tracked by genomic PCR using the T-DNA-specific left border primer (Lba1) in combination with gene-specific primers (see supplemental Table

S1 for all primers used in this study). The exact T-DNA insertion positions were determined by sequencing the T-DNA-specific PCR products using BigDye Terminator Sequencing v3.1 (University of Wisconsin - Madison Biotechnology Center). For RT-PCR, total RNA was extracted from 7-day old liquid grown seedlings of the indicated genotypes using the RNAeasy plant mini kit (Qiagen), and then converted to cDNA using oligo-(dT)₂₀ primers and the SuperScript III first-strand synthesis system (Invitrogen) both as according to manufactures instructions. RT-PCR of the converted cDNA was then performed with primer pairs (P1-P9).

Transgenic plants expressing FLAG-tagged variants of *RPT4a* and *RPT4b* were constructed as follows. The genomic region encompassing the full coding sequence of *RPT4a* (*AT5G43010*), the 600bp sequence upstream of the ATG start codon, which included the 5'-untranslated region (UTR), and the 3'-UTR, was PCR-amplified from Col-0 genomic DNA using primer pair (P10 and P11) and cloned into the pDONR221 (Invitrogen) vector via the Gateway BP Clonase II reaction (Invitrogen). Codons for tandem FLAG (DYKDDDDK) tags were inserted using two rounds of site directed mutagenesis (Edelheit *et al.*, 2009) with primer pairs P12, P13 and P14, P15. The sequence-confirmed 2xFLAG-*RPT4a* clone was recombined into the plant transformation vector pMDC123 (which encodes resistance to the herbicide Basta (glufosinate)) via the Gateway LR Clonase II reaction (Invitrogen). The genomic

region including the full coding sequence of *RPT4b* (*AT1G45000*) (from start to stop codon, lacking the 3UTR), plus the 2kb sequence upstream of the ATG start codon (including 5' UTR) was PCR-amplified from *A. thaliana* Col-0 genomic DNA using the primer pairs P16, P17 (containing the 2XFLAG tag, *NcoI*, and *SmaI* cleavage sites) and P18, P19 (containing a *PstI* and *NcoI* sites) for the coding sequence and upstream region, respectively. The amplified regions and the recipient pCAMBIA3301 vector were digested with *NcoI*, *SmaI*, and *PstI* (New England Biolabs (NEB)), and ligated together using T4 DNA ligase (NEB). The recombinant plasmid was introduced into the *Escherichia coli* TOP10 strain (Invitrogen), and confirmed as correct by sequencing.

The *PAP1* coding sequence was PCR-amplified from Col-0 cDNA using primer pair P20, P21, and cloned into the pDONR221 vector via the Gateway BP Clonase II reaction. The *PAP1* coding sequence was then recombined in-frame into the plant transformation vector pMDC99 which harbored a 3x-HA tag upstream of the recombination site (resulting in a 3x-HA tag, followed by a short linker RSSRGVHHMITT-TLYTKVEMK, followed by the initiator methionine for PAP1) driven by a *UBQ10* promoter. This was accompanied by a hygromycin B resistance gene. All constructs were transformed into plants by the *Agrobacterium tumefaciens*-mediated floral-dip method with the GV3101 agrobacterium strain (Clough and Bent, 1998; Gelvin,

2003; Zhang *et al.*, 2006). The construction encoding *RPT4a::FLAG-RPT4a* and *RPT4b::FLAG-RPT4b* were transformed into plants harboring the *rpt4a-1* and *rpt4b-2* insertions, while the construction encoding *UBQ10::HA-PAP1* was transformed into *PAG1-FLAG pag1-1* plants. T1 plants were selected with the relevant plant resistance markers (Basta for both the *RPT4a* pCAMBIA3301 vector and the *RPT4b* pMDC123 vector, and hygromycin for the *PAP1* pMDC99 vector). T2, and T3 plants were generated by self-pollination. T3 progeny analysis of herbicide/antibiotic-resistant segregation was used to identify homozygous T2 plants. Expression of the FLAG-RPT4a/b transgenes was confirmed by SDS-PAGE followed by immunoblot analysis of total cell extracts with either the M2 anti-FLAG antibodies (Sigma Catalogue Number (Cat. No.) F1804) or anti-RPT4 antibodies generated by Kwang-Hee Lee (University of Wisconsin). Total cell extracts were prepared from plants grown for 7-days on 0.8% agar plates containing standard plant growth medium (GM medium, containing MS salts, and Gamborgs B5 vitamins (Gamborg *et al.*, 1968; Julio and Jose, 2006; Murashige and Skoog, 1962)), and 2% sucrose). Expression of HA-PAP1 was confirmed by SDS-PAGE of total extracts (prepared as above) followed by immunoblot analysis with anti-HA antibodies (Sigma anti-HA antibody, Cat. No. H6908).

1.4.2. Sequence and Phylogenetic Analyses

Plant protein sequences related to *Arabidopsis* PBAC1-4 and PAP1 were obtained using the Basic Local Alignment Search Tool (BLAST) against plant genomes (*Zea mays*, *Oryza sativa*, *Medicago truncatula*, and *Physcomitrella patens*) from Phytozome v8 (Goodstein *et al.*, 2012). Sequences for Pba1-4 (*Saccharomyces cerevisiae*) and PAC1-4 (*Homo sapiens*) were obtained from literature searches. The plant PBAC nomenclature is used for consistency. Related animal, and yeast sequences were obtained using top hits in each respective proteome from iterative PSI-BLAST searches using the BLOSUM62 substitution matrix, and hits above the default expectation value (E-Value) threshold of 0.005 in each iteration (Altschul *et al.*, 1997). Sequence alignments and identity matrices were determined using the Clustal Omega, which selects the appropriate Gonnet 40, 80, 120, 160, 250 and 350 protein substitution matrices in its alignment algorithm (Gonnet *et al.*, 1992; Sievers and Higgins, 2014; Sievers *et al.*, 2011). Alignments were visualized with BOXSHADE (v3.2.3, <https://sourceforge.net/projects/boxshade/>). Phylogenetic trees were generated with MrBayes (v3.2.2) using 1 million generations and a relative burn-in of 25% (Ronquist *et al.*, 2012). The analysis was run until convergence, and the resulting consensus tree was visualized in FigTree (v1.4.2, <http://tree.bio.ed.ac.uk/software/figtree/>).

1.4.3. Proteasome Affinity Purifications using PAG1-FLAG and FLAG-RPT4a/b

26S proteasomes and various subcomplexes were affinity purified based on integrated FLAG epitopes as described previously with a minor modification in the primary buffering agent, which was switched from Tris to HEPES, as it has a more stable pK_a (Book *et al.*, 2010; Marshall *et al.*, 2017). Plant seedlings were grown for 10 days under constant 150 μmol/m²/s light in liquid GM medium containing 2% sucrose at 21-23°C with gentle shaking (90 rpm). For MG132 treatments the growth medium was replaced at day 9 with fresh medium containing 50 μM MG132 (Sell-eckChem) prepared from a 50 mM MG132 solution dissolved in DMSO, or fresh medium containing an equivalent volume of DMSO, and harvested after 16 hours. Frozen seedlings (2.5 - 5g) were pulverized by mortar and pestle at liquid nitrogen temperatures, extracted at 4°C with 1.25 mL/g of fresh weight of buffer A.1 (50 mM HEPES (pH7.5), 50 mM NaCl, 10 mM MgCl₂, and 10% (v/v) glycerol, 2 μM chymostatin, 2mM phenylmethane sulfonyl fluoride (PMSF), 5mM dithiothreitol (DTT), with or without 20 mM ATP), filtered through Miracloth (EMD Millipore), and clarified at 30,000 × g for 30 min. Clarified protein extracts were applied to 50 μL of M2 anti-flag affinity resin (Sigma, Cat. No. A2220) (pre-equilibrated by washing with 1 mL buffer A.1 three times), washed three times with 40 column volumes (2 mL) of buffer A.1. Proteasomes, or their sub-complexes were eluted

in approximately 250 μ L of 500 ng / μ L FLAG peptide (DYKDDDDK synthesized at the University of Wisconsin Biotech Center, Madison WI) in A.1 buffer without PMSF and chymostatin. Samples were stored at -80°C.

1.4.4. SDS-PAGE and Native-PAGE Analysis and Silver Staining

SDS-PAGE utilized a Tris-glycine buffering system and was prepared according to the protocol originally established by Laemmli (Laemmli, 1970) and were performed as described in (Marshall *et al.*, 2017). Gels were made using 40% (w/v) acrylamide/bis-acrylamide solution (29:1, Bio-Rad), SDS-PAGE resolving buffer (375 mM Tris-HCl (pH 8.8) and 0.125% SDS), SDS-PAGE stacking buffer (125 mM Tris-HCl (pH 6.8) and 0.125% SDS), 0.1% ammonium persulphate (APS) and polymerized with N,N,N',N'-tetramethylethylenediamine (TEMED) with a final acrylamide concentration of 11% and 3.5% for the resolving and stacking gels, respectively. The gels were cast using a Hoeffer Gel casting system into 18 x 16 cm glass plates using 0.75 mm spacers. Proteins were separated with a Hoeffer SE600 electrophoresis unit using Tris-glycine SDS running buffer (25 mM Tris (pH 8.3), 192 mM glycine, 0.1% SDS; Bio-Rad) at a constant current of 15 mA for 6 hours. Approximately 32 μ L out of 250 μ L of each affinity-purified proteasome preparation was loaded for SDS-PAGE with Laemmli sample buffer. Native-PAGE gels were

prepared according to Book *et al.* (Book *et al.*, 2010), and performed exactly as described in (Marshall *et al.*, 2017). A discontinuous Native-PAGE system was used. The lower gel contained 4.6% acrylamide (w/v), 0.12% Bis (w/v), 2.3% sucrose (w/v), Tris-Borate-EDTA (TBE) buffer (89 mM Tris-HCl (pH 8.3), 89 mM boric acid, 2 mM EDTA, Bio-Rad), 0.1% APS, 5 mM MgCl₂, 1 mM ATP. The upper gel contained 2.5% acrylamide (w/v), 0.62% Bis (w/v), 2.3% sucrose, TBE, 0.1% APS, 5 mM MgCl₂, 1mM ATP. Gels were cast and run as described above. Gels were run using a Hoeffer SE600 electrophoresis unit with TBE buffer supplemented with 0.5 mM ATP instead of Tris-glycine SDS running buffer. Approximately 32 µL of each proteasome affinity purification was supplemented with 5 µL of 0.03% (w/v) xylene cyanol to visualize loading. Gels were run at 4°C at 50V for 16 or 32 (extended Native-PAGE) hours. Silver-staining of all gels was accomplished using a MS-safe short silver nitrate staining protocol (Chevallet *et al.*, 2006).

1.4.5. MS/MS Sample Preparation

All chemicals for MS/MS analyses were obtained from Sigma-Aldrich, unless otherwise noted. All solvents used were obtained from Thermo-Fisher and were LC/MS grade or J.T Baker® grade unless otherwise noted. For in-solution digests, 200 µL of eluent from each proteasome affinity purification was dried down in a

Speed Vac centrifugal evaporator (Savant Instruments, model number SUC100H) until approximately 25 µL of volume remained, then resuspended in 100 µL of 8 M urea. The samples were reduced with 10 mM DTT (1 hour RT), alkylated with 40 mM iodoacetamide (1 hour RT in the dark), quenched with 40 mM DTT (10 min RT) and then digested with approximately 1 µg of trypsin overnight at 37°C. The samples were acidified with 10% trifluoroacetic acid (TFA) until the pH was 2, desalted on C18 Bond Elut OMIX Tips (Agilent) as according to manufacturers instructions, and eluted in 75% acetonitrile (ACN), 0.1% acetic acid. The samples were then dried in a Speed Vac centrifugal evaporator and resuspended in 30 µL of 5% acetonitrile, 0.1% formic acid and stored at -80°C until analyzed by MS/MS.

For digestion of samples run on a native-PAGE gel, silver-stained protein bands were cut into 1-mm³ pieces with a clean razorblade, washed with MilliQ (EMD Millipore) water, and destained with a 1:1 solution of freshly prepared 100 mM sodium thiosulfate ($\text{Na}_2\text{S}_2\text{O}_3$) and 30 mM potassium ferricyanide ($\text{K}_3\text{Fe}(\text{CN})_6$) until colorless. Silver-ions were then removed from the gel slices by washing twice with MilliQ water. The slices were dehydrated with a 50% ACN in 25 mM ammonium bicarbonate (NH_4HCO_3) solution, then dried in a Speed Vac. They were rehydrated with a reducing solution (25 mM DTT in 50 mM ammonium bicarbonate) for 20 minutes at 56°C, to cleave disulfide bonds. The gel pieces were then treated with

an alkylating solution (55 mM iodoacetamide in 50 mM ammonium bicarbonate) in the dark for 20 minutes at RT to alkylate free cysteines. The alkylating solution was then removed, the gel slices were washed with MilliQ water, dehydrated in 50% ACN in 25 mM ammonium bicarbonate solution, and then dried in a Speed Vac. The ProteaseMax (Promega) in-gel digest protocol was then used to aide in digestion and peptide extraction from gel slices (Saveliev *et al.*, 2013). Samples were then desalted with self-packed C18 stage-tips (Rappsilber *et al.*, 2003) using a spin protocol (Yu *et al.*, 2014), dried, and resuspended as described above.

1.4.6. MS/MS Analyses

In-solution-digested peptide mixtures were separated using a Dionex U3000 nano-flow ultra-high performance liquid chromatography system (nUHPLC, Thermo Scientific). The peptides were separated using an Acclaim® PepMap RSLC C18 column, 2 μ m particle size, 100 \AA pore size, 75 μ m x 15 cm (Thermo Scientific) using a 90 minute linear gradient from 5% ACN to 33% ACN with 0.1% formic acid. Electro-spray ionization (ESI) of separated peptides was performed with a NanoSpray Flex ion source and ionized peptides were subsequently analyzed with a Thermo Q-Exactive Plus high-resolution accurate-mass orbitrap MS/MS. A data-dependent acquisition program was run with an MS1 mass range of 380-1500 m/z

at 70,000 mass resolution, with an automatic gain control (AGC) target of 1e6 ion intensity. The top 15 peaks were chosen for higher energy collisional dissociation (HCD) fragmentation at 28% normalized collision energy with a dynamic exclusion set to 10 s. Product MS2 scans were acquired with a 200-2000 m/z mass range at 17,500 mass resolution, with an AGC target of 8e3. All peaks were recorded in profile mode. All raw files associated with these experiments will be deposited in the PRoteomics IDEntification (PRIDE) database.

Peptide samples obtained from native-gel slices were instead analyzed using a nUHPLC (nanoAcuity, Waters Corporation) connected online to an electrospray ionization LTQ Velos Orbitrap mass spectrometer (Thermo Scientific). Separation of the peptide mixture employed a 100 x 365 μm fused silica capillary micro-column packed with 20 cm of 1.7 μm -diameter, 130- \AA pore size, C18 beads (Waters BEH), with an emitter tip pulled to approximately 1 μm using a laser puller (Sutter Instruments). Peptides were separated with a 2% to 30% acetonitrile gradient in 0.1% formic acid. Full-mass scans were performed in the FT Orbitrap with a mass range of 300-1500 m/z at a resolution of 60,000, followed by ten MS/MS high energy C-trap dissociation scans of the ten highest intensity parent ions at 42% normalized collision energy and 7,500 resolution, with a mass range starting at 100 m/z. Dynamic exclusion was enabled with a repeat count of two over the duration of 30

s and an exclusion window of 120 s.

1.4.7. MS/MS Data Processing

For the proteasome associated protein studies, raw spectra files were processed with MaxQuant version 1.5.3.30 (Cox and Mann, 2008) using the Andromeda search engine against a TAIR10_pep_20101214_updated.fasta file obtained from The *Arabidopsis* Information Resource (TAIR) version 10 (Berardini *et al.*, 2015).

Andromeda searches were performed with MS1 mass tolerance of 20 ppm and an MS2 mass tolerance of 10 ppm. Up to two missed cleavages were allowed, along with the fixed modifications of carbamidomethyl on cysteine, and variable modifications including oxidation of methionine, acetylation of protein N-termini, and GlyGly (ubiquitylation remnant after cleavage with trypsin) on lysines for possible ubiquitylation events, were specified. Peptides and protein groups were identified with a 0.01 false discovery rate (FDR).

Label-free-quantitative analyses were performed using MaxQuants MaxLFQ (Cox *et al.*, 2014) algorithm (generating the MaxLFQ values output) with default settings except that only unique peptides were used for quantification so that protein isoforms could be discriminated. Matching was set to match from and to samples to increase the number of quantifiable proteins per run, using the 'MaxQuants match

between runs' feature. MaxQuants proteinGroups.txt file was then processed in Perseus (Tyanova *et al.*, 2016), a software suite for analyzing mass spectrometry data. Ten proteins were marked as likely contaminants including nitrilases, cruciferins, and seed storage albumins (full list available in Table S2). MaxLFQ values were averaged for both technical replicates, and missing MaxLFQ quantification values were imputed with a 1.8 fold downshifted normal distribution with width 0.3. Volcano plots were generated comparing each proteasome affinity purification against its respective wild-type control using Perseus's volcano plot function using a two-sided t-test, 0.01 FDR, S0 value of 2, and 250 permutations. The resulting volcano plots were plotted using the Seaborn python graphing library (Tyanova *et al.*, 2016; Waskom, 2016), and edited in Adobe Illustrator. Proteins that were called by Perseus as statistically significant were included in tables (x through x). KDE estimations for CP and RP subunits were performed using the Seaborn graphing library, with the median test and Kolmogorov-Smirnov (KS) tests for enrichment performed using `scipy.stats`. For the bar graphs in Figure 3, missing values were imputed with a $\text{Log}_2(\text{MaxLFQ value})$ of 15 so that a "zero value" could be more easily viewed in log scale.

For the MS/MS analyses of native gel slices the data were searched with the Morpheus search engine (Wenger and Coon, 2013) against same protein database

used above. The recommended settings of MS1 search tolerance of 2.1 Da, MS2 search tolerance of 0.01 Da, and 1% maximum FDR were used. Fixed modifications of carbamidomethyl on cysteine, and variable modifications of oxidation on methionine, and GlyGly on lysine with up to two missed cleavages were used. LFQ was completed with Morpheus Spectral Counter (Gemperline *et al.*, 2016) to calculate dNSAF values as described in Chapter 2. The data were plotted in R using the heatmap.2 function, with centroided hierarchical clustering based on Pearsons correlation as the distance function.

1.4.8. Yeast-two hybrid, bimolecular fluorescent complementation, and HA-PAP1 immunopurification

Assays for protein-protein interactions between PBAC1-4 and PAP1 were performed using the ProQuest® two-hybrid system which uses the pDEST22 and pDEST32 vectors, and the *Saccharomyces cerevisiae* MaV203 strain (Vidal *et al.*, 1996) all available from Thermo Fisher Scientific. PBAC1-4 and PAP1 were amplified from *Arabidopsis* cDNA with primers (P20 P29) and transformed into the pDONR221 entry vector (Thermo Fisher Scientific). Genes were then recombined using Gateway LR Clonase II (Thermo Fisher Scientific) into vectors pDEST22 and pDEST32. Yeast strain MaV203 was transformed with pairwise gene combinations in pDEST22

and pDEST32, or the empty vector controls. Interactions were tested by growing transformants for 2 days at 30°C on synthetic complete medium lacking leucine, tryptophan, and histidine, with added 30mM 3-amino-1,2,4-triazole.

For BiFC assays, PBAC1-3 and PAP1 cDNA sequences in pDONR221 were recombined using Gateway LR Clonase II (Thermo Fisher Scientific) into either pSITE-N-EYFP-C1 or pSITE-C-EYFP-C1 vectors (Martin *et al.*, 2009) (ABRC stocks CD3-1648 for NYFP or CD3-1649 for cYFP respectively), and subsequently transformed into *Agrobacterium tumefaciens* strain GV3101 (Gelvin, 2003). Overnight cultures were diluted in resuspension buffer (10 mM MgCl₂, 10 mM MES (pH 5.7)) to a final OD₆₀₀ of 0.5 and then infiltrated by syringe into 4-6 week old *Nicotiana benthamiana* leaves. After 36 h, fluorescence in the infiltrated regions was visualized using a Zeiss 510 Meta confocal laser scanning microscope, and images were analyzed with Zeiss LSM image browser v3.5 and ImageJ.

The immunopurification of PAP1 to determine if PAP1 could co-purify with 26S proteasome subunits was performed as follows. HA-PAP1-expressing tissue in the *PAG1-FLAG pag1-1* background were grown in liquid culture as described previously (Book *et al.*, 2010). Approximately 2 g of tissue were ground using liquid nitrogen chilled mortar and pestles, and soluble proteins were extracted in 2.5 mL of buffer A.1. 25 µL of HA-EZview (Sigma-Aldrich) beads were incubated overnight

with 1.75 mL of protein extracts in 2 mL Eppendorf tubes at 4°C with constant rotation. The HA beads were washed with 3 x 1mL of buffer A.1, and then eluted with 100µL of 2x protein sample buffer ((Laemmli, 1970)).

1.5. Results

1.5.1. *Arabidopsis* Proteasomes Can Be Effectively Affinity-Purified Using Epitope-Tagged Versions of Either Isoform of the Regulatory Particle Subunit RPT4

As a first step toward purifying *Arabidopsis* 26S proteasome via the RP, and exploring the possibility that plants assemble unique proteasome isotypes, we sought out candidate subunits amenable to epitope tagging for which useful mutants were available for genetic replacement. One of the best candidates was *RPT4*, which is encoded by two isoforms (*RPT4a* and *RPT4b*) with 96% amino-acid sequence identity. This level of identity is near average for shared sequence identity among RPT subunits (Book *et al.*, 2010); RPT2a and RPT2b share the most identity at 99%, and RPT5a and RPT5b share the least identity at 93%. Based on the structure of the yeast 26S proteasome (PDB ID: 4CR2) (Beck *et al.*, 2012), we expected that the N-terminus of RPT4 would be solvent accessible, thus making it an attractive candidate for attaching an tag that would be exposed and minimally disrupt particle

assembly. Furthermore, we identified several T-DNA insertion mutations within *RPT4a* and *RPT4b* in the Col-0 background that might generate null alleles suitable for genetic replacement (Figure 1.1 A). DNA sequence analysis of the *rpt4a/b* mutants determined the insertion sites to be at nucleotides 681, -278, and -24 for *rpt4a-1*, *rpt4a-3* and *rpt4a-4* and 151, 409, and 558 for *rpt4b-1*, *rpt4b-2*, and *rpt4b-3*, from the ATG start codon respectively. The *rpt4a-1* and *rpt4b-2* alleles were particularly notable because the T-DNA insertions sites disrupted the coding regions at 88 and 101 residues upstream of the AAA-ATPase cassette (Figure 3.S1) for each polypeptide which would likely block accumulation of the full length transcript (Figure 1.1 B). This possibility was confirmed by RT-PCR analysis of RNA extracted from seedlings homozygous for either *rpt4a-1* or *rpt4b-2* showing that the plants failed to accumulate the corresponding full-length transcripts (Figure 1.1 B, Primers P1 + P5, and Primers P6 + P9 respectively). However, low levels of the 5' region was amplified using primers that were 5' to the insertion site in both *rpt4a-1* and *rpt4b-2* plants (Primers P1 + P2, and Primers P6 + P7) suggesting that a transcript for the 5' region still accumulates. Additionally, a small amount of product relative to wild type was also detected 3' to the insertion site for *rpt4a-1* suggesting some mRNA from this region can also accumulate. At the protein level, an immunoblot analysis of SDS-PAGE-separated proteins from *rpt4a-1*, *rpt4b-2* and wild type Col-0

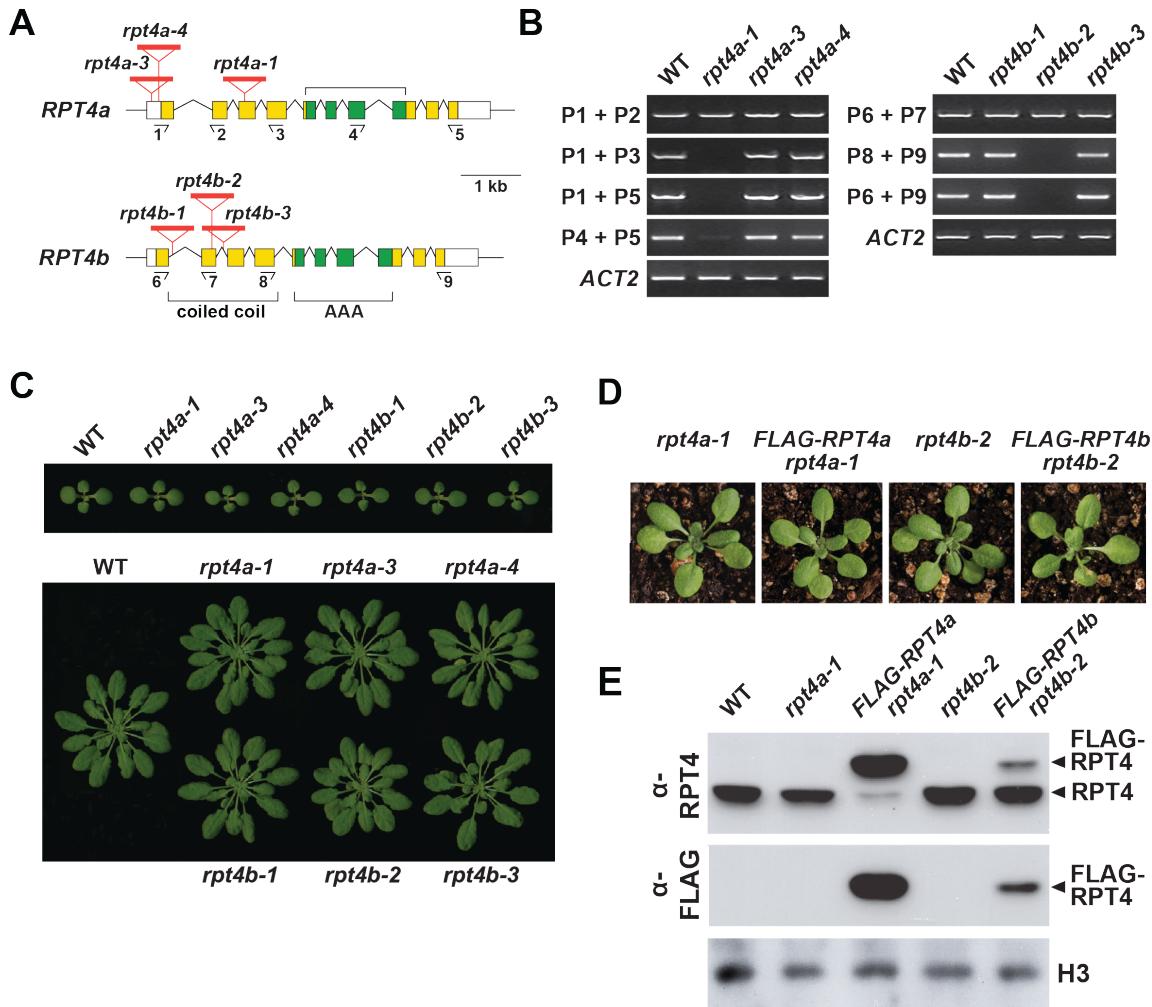


Figure 1.1: *rpt4a-1* and *rpt4b-2* mutants rescued with flag tagged variants. (A) Organization of the *Arabidopsis* *RPT4a* and *RPT4b* genes with insertion points shown for T-DNA alleles of *RPT4a* and *RPT4b*. **(B)** RT-PCR analysis of the *RPT4a* and *RPT4b* transcripts in wild type and *rpt4a* and *rpt4b* mutant seedlings showing aberrant *RPT4a* transcripts in *rpt4a-1* and abnormal *RPT4b* mRNAs in *rpt4b-2* plants. **(C)** Homozygous *rpt4* mutants show a phenotype similar to wild-type control. The top panel shows plants grown for 10-days in a 16-hour photoperiod, while the lower panel shows plants grown for 42 days in an 8-hour photoperiod. **(D)** Replacing the wild-type *RPT4a* and *RPT4b* with their respective flag-tagged variants show plants with normal growth when grown for 7 days on GM plates and then grown in soil for two weeks. **(E)** Immunoblots using anti-RPT4 and anti-FLAG antibodies confirm the presence of the tag and an increased size due to the FLAG tag, which is reactive to the anti-FLAG antibody. A loading control for histone-3 is shown.

seedlings showed reduction in RPT4 protein in *rpt4a-1* compared to wild type, but little to no change in *rpt4b-2* (Figure 1.1 E). While RPT4a is the dominant isoform with 169 expressed sequence tags as compared to 39 for RPT4b (Berardini *et al.*, 2015), our data detected a stronger reduction in RPT4a levels, as compared to RPT4b. There were no differences in RPT4 immunoreactivity between RPT4a and RPT4b (communications with Kwang-Hee Lee), therefore a more complex regulation of global RPT4a/b expression than anticipated may occur. The mutants were phenotypically indistinguishable to wild-type when grown on agar plates for 10 days in a long day (LD) photoperiod (16 h light / 8 h dark) or when transferred to unfertilized soil and grown until 42 days old in a short-day photoperiod (8 h light / 16 h dark) (Figure 1.1 C). Taken together these data suggest that the two RPT4 isoforms are not essential by themselves.

Plants containing the mutant alleles *rpt4a-1* and *rpt4b-2* were transformed with their respective 2XFLAG-tagged genomic variants and selected for stable transgene insertion. Each construct included codons for a 2XFLAG tag and the genomic region including introns that bracketed the *RPT4a/b* coding regions with expression drive by the promoter and 5' UTR regions of each locus (see Experimental Procedures). The initial transformants were self-fertilized and the homozygous lines for both the *rpt4* T-DNA insertion and the transgene were identified in the Basta

resistant progeny by genomic PCR. Immunoblot analysis of proteins extracted from *RPT4a/b::FLAG-RPT4a/b*-transformed plants detected accumulation of the FLAG-RPT4 proteins, which were larger than the corresponding native RPT4 polypeptides when detected with anti-RTP4 antibodies, and could be detected with anti-FLAG antibodies (Figure 1.1 E). The lower bands in the anti RPT4 immunoblot were likely RPT4b in the case of the *rpt4a-1* mutant, and RPT4a in the case of the *rpt4b-2* mutant. The transformed plants were phenotypically indistinguishable from wild type, implying that ectopic expression of either isoform bearing the 2X-FLAG tag was not detrimental to growth and development (Figure 1.1 D). Taken together these data showed that transgenic plants expressing their respective FLAG-tagged RPT4 variants displayed no gross morphological defects.

Using the strategy first described by Book *et al.* (Book *et al.*, 2010), we exploited the homozygous *FLAG-RPT4a/b rpt4a/b* seedlings to affinity purify 26S proteasomes via the RP. Briefly, frozen seedlings were homogenized in a tissue extraction buffer (see Experimental Procedures), clarified by centrifugation, and subjected to anti-FLAG chromatography, using FLAG peptide for gentle elution. To enrich for the 26S particle, 20 mM ATP was included in the indicated buffers to stabilize the CP/RP association (Book *et al.*, 2010; Liu *et al.*, 2006). To enrich for just the RP (using FLAG-RPT4a/b) or the CP (using PAG1-FLAG) the same protocol was used

except that ATP was omitted from the buffers. As a control, we also attempted to purify 26S proteasomes from wild-type tissues. Under conditions where 20 mM ATP was included in the purification buffer (indicated with a + in Figure 1.2 A), the FLAG-RPT4a/b affinity purified samples showed a comparable protein banding pattern after SDS-PAGE as compared to the PAG1-FLAG affinity-purified samples suggesting a similar complement of proteins were present (Figure 1.2 A). The only notable exception was the observation of a band that likely represented the FLAG-tagged PAG1 subunit (indicated with a black arrow) in the PAG1-FLAG samples (Book *et al.*, 2010). The size-shifted RPT4 subunits seemed to be masked by other subunits and were not readily identifiable via silver staining alone. Importantly, the mock affinity-purified samples from wild-type tissue showed very few contaminating proteins as judged by protein staining with silver, which was important for down-stream tandem mass-spectrometry (MS/MS) analyses. Overall, these affinity purification strategies based on either RPT4a or RPT4b successfully purified 26S proteasomes from whole *Arabidopsis* seedlings. When ATP was omitted from the purification buffer, there was clear enrichment for the CP in PAG1-FLAG samples, and clear enrichment for the RP in both the FLAG-RPT4a and FLAG-RPT4b samples, as judged by the size and intensity of the protein bands in Figure 1.2 A. Immunoblot analysis of these samples showed clear enrichment for two CP subunits PAG1 and

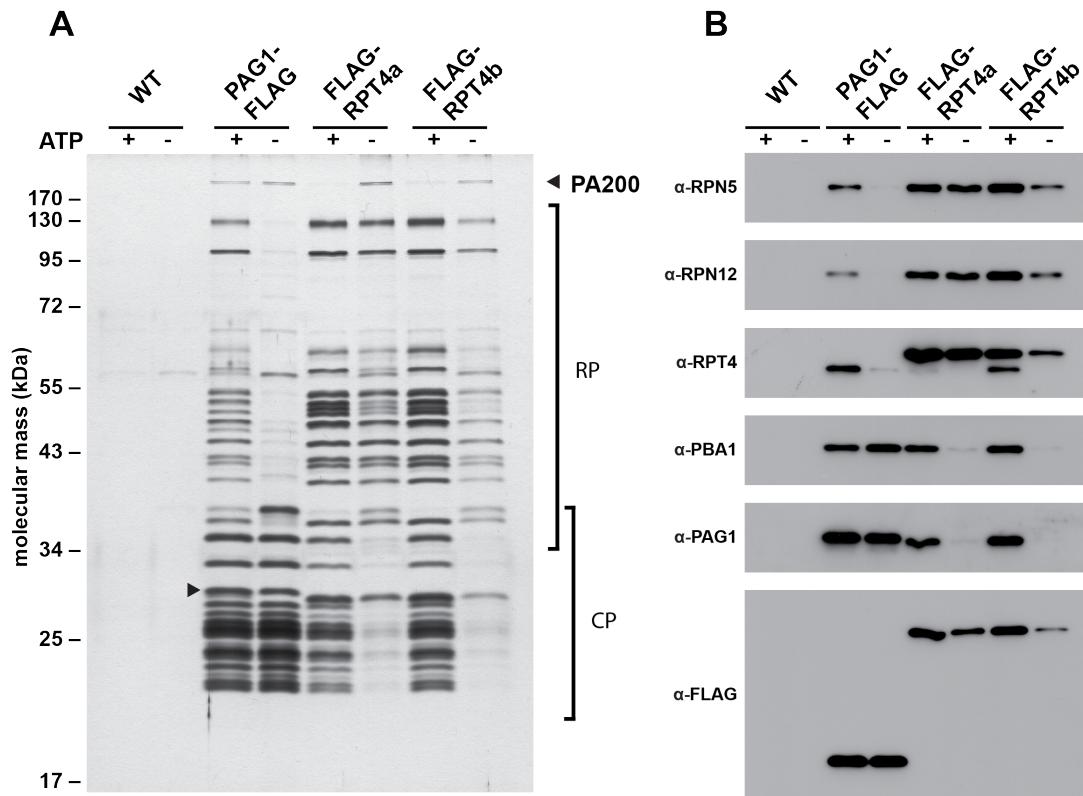


Figure 1.2: Protein composition of affinity-purified proteasomes from FLAG-RPT4a-, FLAG-RPT4b-, and PAG1-FLAG-expressing plants and from mock wild-type controls in the presence (+) or absence (-) of 20mM ATP. (A) Silver stained gel of each preparation shows clear enrichment over the negative control (Col-0) with minimal contaminants, both in the presence and absence of ATP during purification. Omitting ATP from the preparations for the CP-based PAG1-FLAG pulldown shows a decreased amount of RP subunits, while both RP-based subunit pulldowns show a decrease in the amount of CP subunits. (B) Immunoblots for various proteasome subunits for both the RP and the CP show a clear enrichment of proteasome subunits as compared to the wild type mock affinity purification control. As shown in the silver stain, omitting ATP from the preparations in each affinity purification causes a decrease in association of the alternative sub-particle (decrease in RP for CP pulldowns, and decrease in CP for RP pulldowns).

PBA1 in the CP-based samples affinity-purified in the absence of ATP, as compared to the RP subunits RPT4, RPN5, and RPN12 (Figure 1.2 B). There was also clear enrichment for the three RP subunits in the RP-based affinity purified samples when ATP was omitted from the purification as compared to the CP subunits that were immunoblotted. Taken together, these data show that 26S proteasomes can be easily purified based on a RP affinity tag, and that both isoforms for a single subunit will work. By omitting ATP from the buffer, we were able to enrich for subunits of the CP or the RP utilizing affinity purifications from PAG1-FLAG and FLAG-RPT4a/b expressing tissues, respectively.

1.5.2. MS/MS Analyses of RPT4a/b Affinity-Purified Proteins Identifies a Similar Complement of Proteins as PAG1-FLAG Affinity Purifications

Previous MS/MS analyses of affinity preparations based on PAG1-FLAG identified the fundamental subunits that make up the 26S proteasome, including the CP α -subunits (PAA-PAG), the CP β -subunits (PBA-PBG), the RP AAA-ATPases (RPT) (RPT1-6), and the RP non-ATPase subunits (RPN) (RPN1-3 and RPN5-12). Most subunits are encoded as a paralogous gene pair (named 1 and 2 for CP subunits, and a and b for RP subunits) and most isoforms were found to be incorporated in the complex (Book *et al.*, 2010). We wanted to verify that our affinity purifications

based on RPT4a and RPT4b also contained a similar set of proteins as in Book *et al.* (Book *et al.*, 2010). To do this we used label-free quantitative tandem mass spectrometry (LFQ-MS/MS) to analyze samples obtained from affinity purifications for FLAG-RPT4a, FLAG-RPT4b, PAG1-FLAG, and a mock wild-type control, in the presence and absence of ATP resulting in MaxLFQ values (Cox *et al.*, 2014) for each protein analyzed. Importantly, we identified all subunit isoforms previously identified by Book *et al.* (Book *et al.*, 2010) showing that our purifications were comparable (Figure 1.3, CP α -ring, CP β -ring, RPT, and RPN plots). Additionally, we identified a suite of key accessory proteins found to be statistically enriched as compared to control at a false discovery rate (FDR) of 0.01 for both the CP and RP (Figure 1.3 Associated Proteins).

The CP α - and β -subunits show higher MaxLFQ values in the PAG1-FLAG samples as compared to either FLAG-RPT4a/b sample under both plus and minus ATP conditions (Figure 1.3, CP α -ring, CP β -ring). This suggests some free CP was purified even in the presence of ATP, and that the complexes purified were not exclusively RP-CP proteasomes. In the PAG1-FLAG samples, the CP α and β -subunits showed higher MaxLFQ values when ATP was omitted from the affinity purification buffer as compared to when ATP was present (Figure 1.3, CP CP α -ring, CP β -ring). This increase is consistent with the preferential purification of the CP

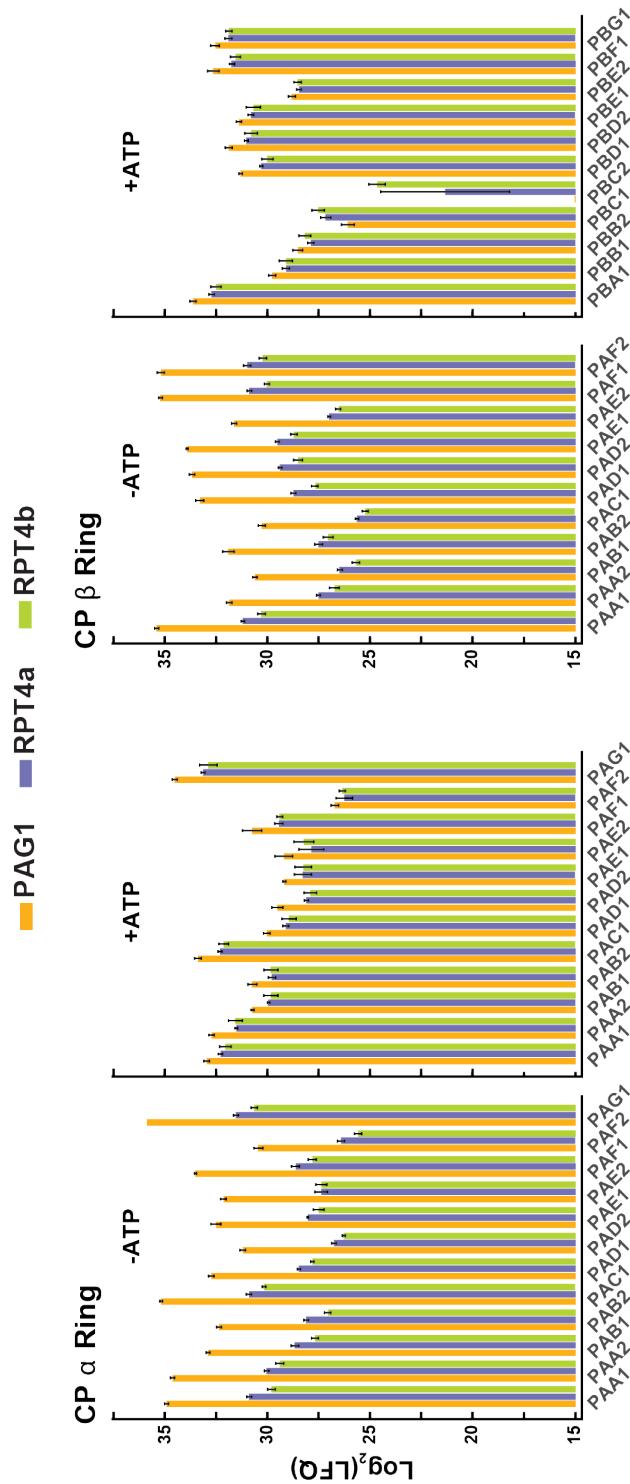


Figure 1.3: Label-free quantitative MS/MS analysis of affinity-purified proteins in the presence (+) or absence (-) of ATP. (CP Alpha Ring) α -subunits are enriched in samples that were affinity purified via the CP subunit PAG1-FLAG. This enrichment increases even further when ATP is omitted from the preparations as shown on the left side of the graph. (CP Beta Ring) β -subunits are enriched at a higher level when purified via the CP subunit PAG1-FLAG. This enrichment increases even further when ATP is omitted from the preparations as shown on the left side of the graph.

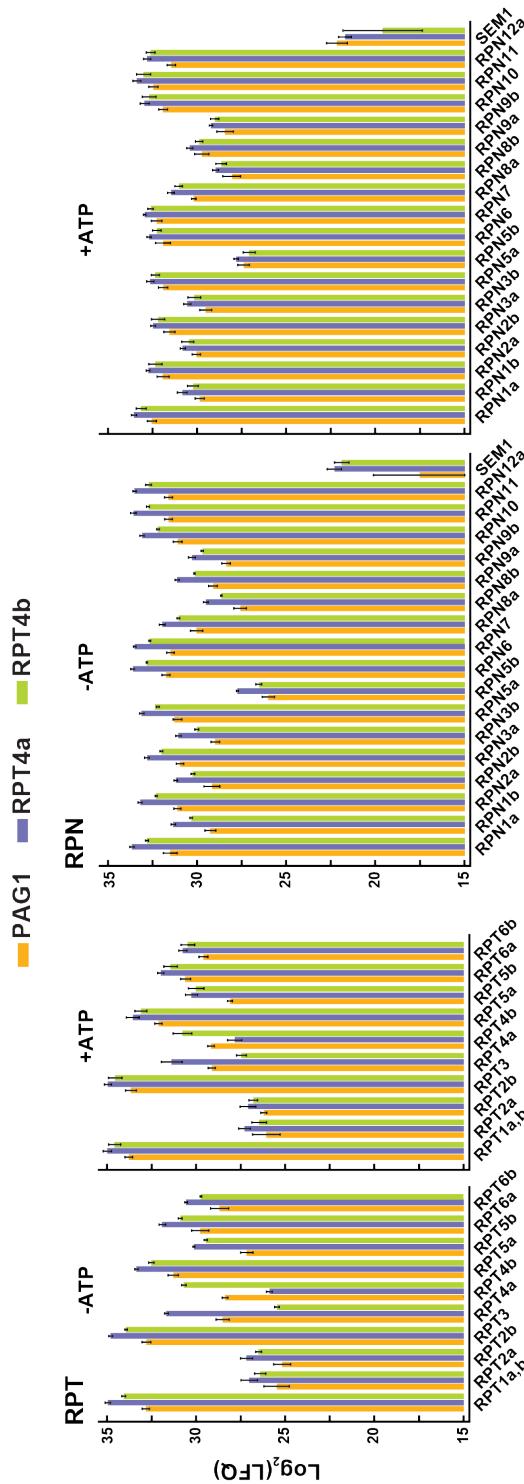


Figure 1.3: Label-free quantitative MS/MS analysis of affinity-purified proteins in the presence (+) or absence (-) of ATP. (RPT Subunits) RPT subunits are enriched at a higher level when purified via RP subunits RPT4a or RPT4b. There are clear increases in the level of enrichment of the RPT4a subunit for the RPT4a affinity purifications, and clear increases in enrichment of the RPT4b subunit for the RPT4b affinity purifications. These respective increases are greater than the isoform unbiased PAG1 pulldowns. The respective enrichments of each isoform also increases when ATP is omitted, due to a decrease in the possibility of binding the alternative subunit isoform under ATP limiting conditions. **(RPN Subunits)** RPN subunits are enriched at a higher level when purified via RP subunits RPT4a and RPT4b. This enrichment increases even further when ATP is omitted from the preparations as shown on the left side of the graph.

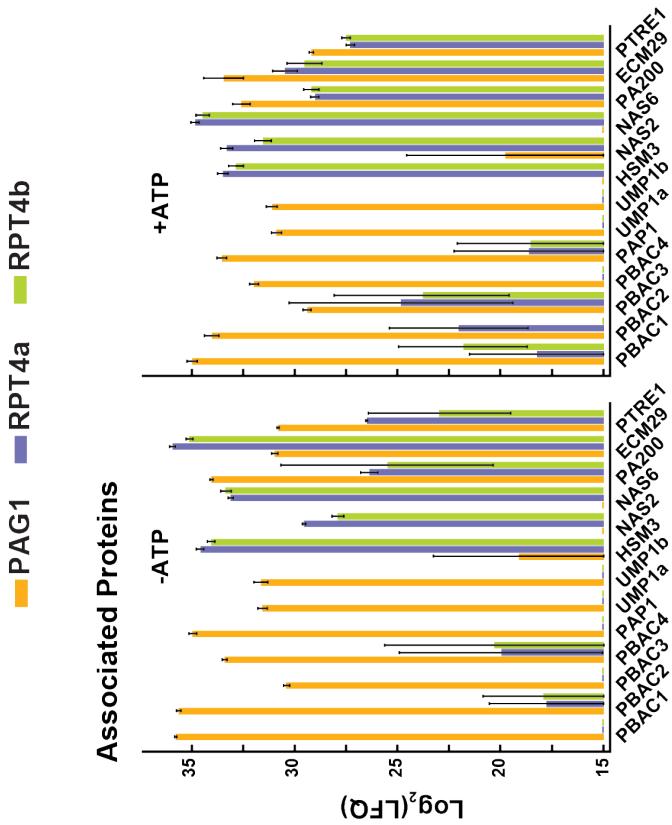


Figure 1.3: Label-free quantitative MS/MS analysis of affinity-purified proteins in the presence (+) or absence (-) of ATP. (Associated Proteins) Key proteins identified as proteasome interacting proteins are specifically enriched in pulldowns based on the CP subunit PAG1, and others are specifically enriched for pulldowns based on either RP subunits RPT4a or RPT4b. Please refer to Figure 4 for statistical analyses of CP and RP distributions in affinity-purified samples, and for statistical analyses that identify candidates as associated proteins.

over the RP in the absence of ATP when using the PAG1-FLAG epitope tag.

Likewise, both affinity purifications that target the RP through FLAG-RPT4a/b, had higher MaxLFQ values for RP subunits as compared to the CP-based PAG1-FLAG samples under both plus and minus ATP conditions. When ATP was omitted from the purification, the FLAG-RPT4a/b samples had increased MaxLFQ values for RP subunits as compared to when ATP was included; however, this increase was considerably less than that observed for the CP subunits in the PAG1-FLAG samples (Figure 1.3, RPT, and RPN plots). FLAG-RPT4a samples had higher MaxLFQ values for RPT4a when compared to FLAG-RPT4b and PAG1-FLAG (Figure 1.3, RPT plot). Conversely, FLAG-RPT4b samples had higher MaxLFQ values for RPT4b when compared to FLAG-RPT4a and PAG1-FLAG (Figure 1.3, RPT plot). These data suggest that, as expected, RPT4a was enriched in the FLAG-RPT4a samples, and that RPT4b was enriched in the FLAG-RPT4b samples.

1.5.3. Statistical Analysis of Affinity Enrichments for FLAG-RPT4a Show Little to No Difference When Compared to FLAG-RPT4b

While mammals and yeast can form only a few types of proteasomes, *Arabidopsis*, with its many subunit paralogs, can potentially form a much wider array of proteasome isotypes. By comparing samples enriched for RPT4a and RPT4b us-

ing a Significant Analysis of Microarray (SAM) test at 0.01 FDR, we determined if any additional subunit isoforms (or other associated proteins) are incorporated preferentially, which would suggest that *Arabidopsis* can form proteasome isotypes. Surprisingly, this statistical analysis revealed that only when ATP was omitted from the purifications were the samples statistically enriched for either RPT4a or RPT4b, respectively (Figure 3.S2 B). While RPT4 and RPT4b were enriched in the samples in which ATP was included in the purification, they were not statistically significant by the SAM test at 0.01 FDR (Figure 3.S2 A). Additionally, the 26S proteasome subunits have larger MaxLFQ values in the FLAG-RPT4a sample suggesting a slightly more efficient purification for the proteasome with the RPT4a isoform. Furthermore, no significant enrichment of any other subunit isoforms was observed. As such these data suggest that there are no major differences in protein composition between affinity purifications for FLAG-RPT4a or FLAG-RPT4b other than statistically enriching for each individual isoform when ATP was omitted from the purification. This indicates that plants may assemble their proteasome randomly with respect to isoform incorporation under the conditions tested.

1.5.4. Statistical Analyses of LFQ-MS/MS Data Show Affinity Preparations Targeting Either the CP or RP Specifically Enrich for Their Respective Sub-complexes

The LFQ-MS/MS data allowed us to perform a more formal statistical test for the enrichment of the RP over the CP in samples affinity purified from the RP subunits FLAG-RPT4a/b, and for the enrichment of the CP over the RP in samples affinity purified from CP subunit PAG1-FLAG. We estimated the distribution of RP and CP subunits in these samples using kernel density estimations (KDE) of MaxLFQ values as compared to our wild-type mock-affinity purifications as a negative control (Figure 1.4 below each volcano plot, and larger plots in Figure 3.S3). As the KDE distributions were not normal, two non-parametric statistical tests, Kolmogorov-Smirnov (KS) and median test, were used to determine if there was specific enrichment of particular proteasome sub-complexes (RP vs CP) from the different affinity purifications. The KS test determines if the two distributions are different in shape or position, and the median test determines if the distributions medians are statistically different.

The KDE distribution for the CP as compared to the RP from samples affinity purified in the absence of ATP using PAG1-FLAG showed both the KS and the median *p*-values being much less than 0.01, at 1.74E-10 and 4.3E-10, respectively

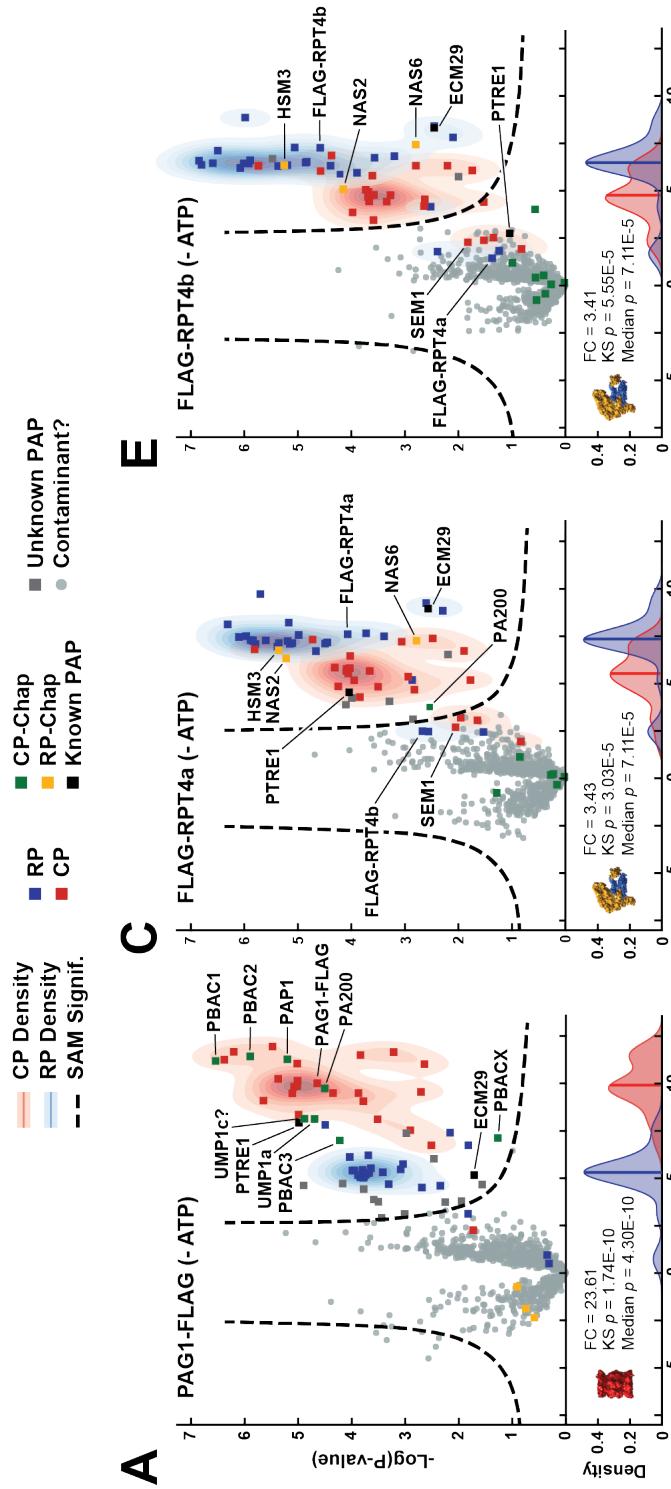


Figure 1.4: Volcano Plots comparing proteasome affinity purifications under both plus and minus ATP conditions relative to WT controls, allow identification of statistically significant interacting partners for the proteasome. Volcano plots comparing affinity purifications for PAG1-FLAG (A and B), FLAG-RPT4a (C and D), and FLAG-RPT4b (E and F) against their respective WT mock affinity purification control are shown. The top panels (A, C, and E) show comparisons for samples purified in the absence of ATP, while the lower panels (B, D, and F) show comparisons for samples purified in the presence of ATP. Two-dimensional Kernel Density Estimations (KDE) are shown for CP subunits in red and RP subunits in blue. Larger KDE plots are available in Figure 3.S3. CP subunits are labeled as red squares, while RP subunits are labeled as blue squares. Proteins outside the dashed lines are statistically significantly differentially represented in purified samples relative to control as judged by a significant analysis of microarray or (SAM) test at 0.01 false discovery rate. Proteins likely involved in CP assembly are shown as green squares, while proteins likely involved in RP assembly are shown as yellow squares. Known proteasome associated proteins or PAPs are labeled as black squares, while unknown PAPs are labeled in grey squares. Proteins identified that did not reach SAM significance are labeled as grey circles.

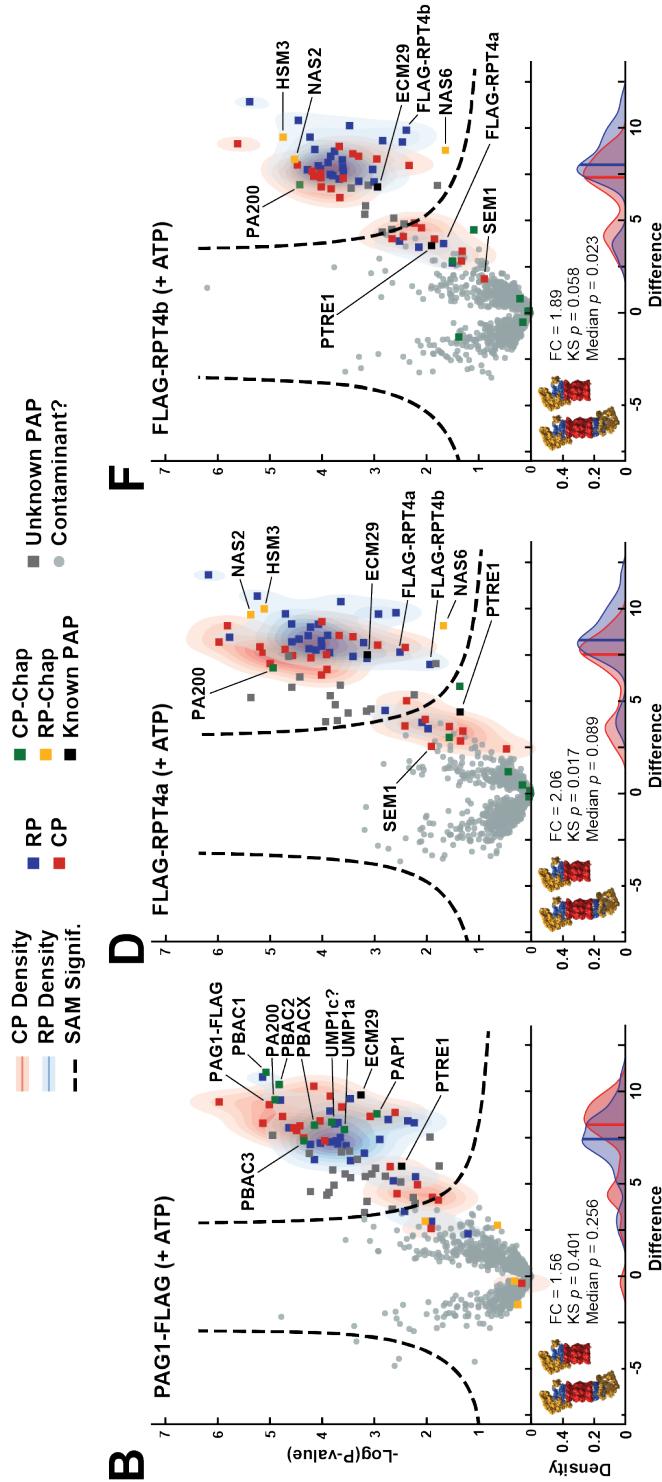


Figure 1.4: Volcano Plots comparing proteasome affinity purifications under both plus and minus ATP conditions relative to WT controls, allowing identification of statistically significant interacting partners for the proteasome. Volcano plots comparing affinity purifications for PAG1-FLAG (A and B), FLAG-RPT4a (C and D), and FLAG-RPT4b (E and F) against their respective WT mock affinity purification control are shown. The top panels (A, C, and E) show comparisons for samples purified in the absence of ATP, while the lower panels (B, D, and F) show comparisons for samples purified in the presence of ATP. Two-dimensional Kernel Density Estimations (KDE) are shown for CP subunits in red and RP subunits in blue. Larger KDE plots are available in Figure 3.S3. CP subunits are labeled as red squares, while RP subunits are labeled as blue squares. Proteins outside the dashed lines are statistically significantly differentially represented in purified samples relative to control as judged by a significant analysis of microarray or (SAM) test at 0.01 false discovery rate. Proteins likely involved in CP assembly are shown as green squares, while proteins likely involved in RP assembly are shown as yellow squares. Known proteasome associated proteins or PAPs are labeled as black squares, while unknown PAPs are labeled in grey squares. Proteins identified that did not reach SAM significance are labeled as grey circles.

(Figure 1.4 and 3.S3 A). This suggests that the distributions are different, and that their medians were different with a greater than 20 fold increase in CP over RP. Conversely the KDE distributions for the samples affinity purified in the presence of ATP using PAG1-FLAG (Figure 1.4 and 3.S3 B) overlap to a higher degree, with both the KS and Median *p*-values > 0.01 at 0.401 and 0.256, respectively. This indicates that in samples affinity purified in the presence of ATP that the RP and CP distributions were not significantly different, and that their means were similar. Although not statistically significant, a 1.5 median fold change was observed for the CP over the RP in samples affinity purified in the presence of ATP using PAG1-FLAG suggesting that there may be some CP enriched as compared to the RP. These statistical analyses showed that, as expected, we were able to successfully enrich for the CP over the RP when affinity purifying from PAG1-FLAG expressing tissue.

Conversely, in samples using the RP-based FLAG-RPT4a/b purified in the absence of ATP, the KDE distributions showed enrichment for the RP over the CP, with a fold change of 3.43 and 3.41 for the RPT4a and RPT4b subunits respectively (Figures 1.4 and Figures 3.S3 C and E, respectively). Both the KS (3.303E-5 for RPT4a and 5.55E-5 for RPT4b) and median *p*-values (7.11E5 for RPT4a and 7.11E-5 for RPT4b) were < 0.01, suggesting that the two distributions were statistically different and had different medians. Again, contrasting this with the samples

purified in the presence of ATP using FLAG-RPT4a and FLAG-RPT4b (Figures 1.4 and Figures 3.S3 D and F, respectively), the KDE distributions overlap with both the KS (0.017 for RPT4a and 0.058 for RPT4b) and median *p*-values (0.089 for RPT4a and 0.023 for RPT4b) were > 0.01 , suggesting that the RP and CP distributions are similar. Although not statistically significant, we still observed a two-fold-change enrichment in both FLAG-RPT4a/b samples purified in the presence of ATP for the RP over the CP. In summary, these KDE distributions and statistical analyses provide further support for enrichment of either the CP or the RP, as expected, when samples were affinity purified in the absence of ATP via PAG1-FLAG, or FLAG-RPT4a/b, respectively.

1.5.5. Statistical Analysis of Affinity Purifications Enriched for the RP or CP Identifies Proteasome-Associated Proteins and Putative Assembly Chaperones Specific to Each Subcomplex

The ability to affinity-enrich both the CP and the RP from *Arabidopsis* under gentle non-denaturing conditions without the need for high salt washes may enable identification of associated proteins, particularly those which may be more loosely bound. Additionally, targeting the RP or CP enables the identification of proteins that may be associating with either or both subcomplexes specifically. The previous

data represented in Figure 1.3 showed the raw MaxLFQ values for some of the relevant associated proteins; however, a formal statistical test against mock affinity purifications was important given the high number of proteins identified and quantified via LFQ-MS/MS. To better understand the suite of proteins that are statistically enriched in samples that target the CP or RP specifically we compared the LFQ-MS/MS data obtained from our affinity purifications against their respective wild-type controls using volcano plots and SAM tests at 0.01 FDR.

Affinity purifications enriching for the CP via PAG1-FLAG, performed either in the presence or absence of ATP (Figure 1.4 A and 1.4 B respectively) show statistical enrichment for two orthologs of the UMP1 CP assembly chaperone, UMP1a and UMP1b. The *Arabidopsis* genome encodes a third *UMP1* ortholog, *UMP1c*; however, we have no experimental evidence that UMP1c associates with the proteasome, even though it has several tryptic peptides from an extended C-terminal domain that could differentiate it from UMP1a/b (Figure S4). UMP1a, and UMP1b are considerably higher expressed at 90 and 55 ESTs as compared to UMP1c which had only 7 ESTs (Berardini *et al.*, 2015), suggesting that UMP1c may be lowly expressed, or possibly a pseudogene.

Intriguingly, these samples also show statistical enrichment for a suite of putative CP assembly chaperones (Proteasome Biogenesis Associated Chaperone)

PBAC1-4 (Figure 1.4 A and 1.4 B). Although their annotation in The *Arabidopsis* Information Resource v10 (TAIR10) was incomplete, an InterProScan of their respective sequences supported these identifications (Table S3A). While PBAC1 was annotated as having a PBAC2 domain, the expectation value of 3.4E-6 was much lower than the expectation value of 1E-29 for the PBAC2 domains in the potential PBAC2 homologue. Additionally, PBAC1 and PBAC2 were previously identified by iterative PSI-BLAST with the human PAC1 and PAC2 orthologs in two separate studies (Kusmierczyk *et al.*, 2011; Le Tallec *et al.*, 2007) which aids in our assignment of PBAC1 and PBAC2. This PSI-BLAST based analysis was repeated using human PAC1-4, which identified PBAC1-4 as the top hits for these proteins in *Arabidopsis* (Table S3B). Protein sequence alignments for PBAC1-4 show that these proteins are conserved across plants but that they show little conservation with orthologs outside the plant kingdom with < 15% protein sequence identity (Figure S5-S8). All statistically significant interactors from PAG1-FLAG affinity purifications performed in both the presence or absence of ATP are available in Table S4.

As shown in Figure 1.4 C-F, UMP1a/b and PBAC1-4 were not significantly enriched in our affinity purifications targeting the RP, suggesting that these proteins interact with the CP specifically. However, affinity purifications that targeted the RP through either FLAG-RPT4a or FLAG-RPT4b showed statistically significant

enrichment over controls for the putative orthologs of RP assembly chaperones HSM3, NAS2, and NAS6 (Figure 1.4 C-F). Importantly, these putative RP assembly chaperones were also absent from our affinity purifications that targeted the CP, suggesting that these proteins are RP specific interactors. All statistically significant interactors from FLAG-RPT4a and FLAG-RPT4b affinity purified in the presence or absence of ATP are available in Table S5 and Table S6. Several accessory proteins are common to both CP and RP purifications including the alternate capping particle PA200 (Book *et al.*, 2010), ECM29 that is involved in stabilizing the CP-RP association (Lehmann *et al.*, 2010), and the recently identified proteasome inhibitor 31 related protein PTRE1, which is involved in modulating proteasome activity in response to auxin (Yang *et al.*, 2016).

1.5.6. A Novel Plant Proteasome-Associated Protein, PAP1, Interacts with a Putative CP Assembly Chaperone

An intriguing plant-specific protein, named here Proteasome Associated Protein 1 or PAP1 was identified and significantly enriched in affinity preparations that targeted the CP (Figure 4). While PAP1 had no domains of known function, as determined by InterProScan (Table S3A), it did contain a C-terminal HbYX (hydrophobic amino acid, followed by a tyrosine, followed by any amino acid) motif. This HbYX motif

is commonly found in proteins that intercalate between the α -rings of the CP (Kusmierczyk *et al.*, 2011). This motif is found in both the plant specific PAP1, and in the assembly chaperone PBAC1 (Figure 1.5 A) as well as the conserved yeast assembly chaperones Pba1 and Pba2 (Kusmierczyk *et al.*, 2011). However, Pba2/PBAC2 orthologs outside of yeast typically contain a C-terminal HbF motif, with the only notable exception being the moss *Physcomitrella patens* which instead contains a HbYX motif. To confirm whether PAP1 did indeed interact with the proteasome, we generated transgenic plants expressing HA-tagged PAP1 in the *PAG1::PAG1-FLAG pag1-1* background. Figure 1.5 B shows immunoblots of SDS-PAGE separated samples from an affinity purification of HA-PAP1 utilizing anti-HA beads. HA-PAP1 expressed at very low levels in transgenic plants (Figure 1.5 B, input). However, upon immunopurification and detection with anti-HA antibodies we were able to detect HA-PAP1, the FLAG-tagged CP α -subunit PAG1, and the FLAG-less PAG1 product that likely resulted from cleavage of the FLAG tag during overnight incubation (Book *et al.*, 2010). Low levels of both the CP β -subunit PBA1, and RPN1 subunit were also detected in these samples.

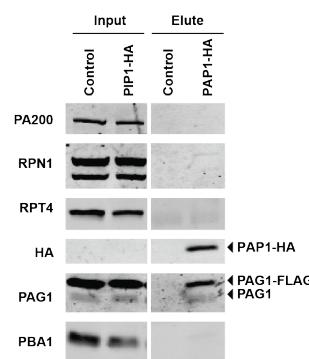
To identify which putative *Arabidopsis* assembly chaperones might interact, we performed a yeast two-hybrid (Y2H) interaction assay for all pairwise combinations of PBAC1-4, and PAP1 (Figure 1.5 C). We hypothesized that PAP1 may be a CP

[Figure 1.5 caption follows on next page]

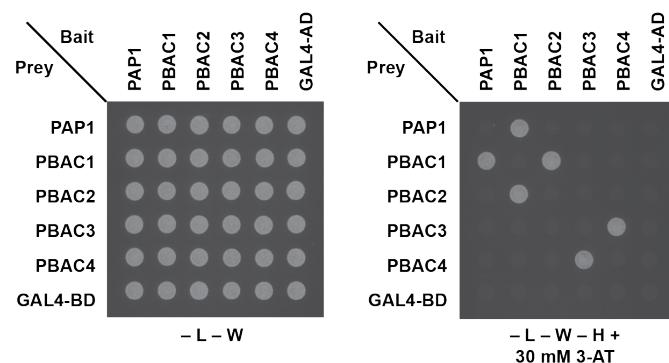
A

	PAP1	PBAC1	PBAC2
<i>A. thaliana</i>	E E S P W R A L Y G	- - L D S D L Y T	P P D M S M F - - -
<i>Z. Mays</i>	D Q E P W R E L Y G	- - D L D S L Y T	P P D M S I F - - -
<i>O. sativa</i>	F Q E P W R E L Y Q	- - E L D A L Y T	P P D L S I F - - -
<i>M. truncatula</i>	I K E P W R D L Y G	- - F Q S H L Y I	P P D V S I F - - -
<i>P. patens</i>	N E D D W R R L Y I	L L S D S E T Y M	P P D D T M F C - - -
<i>S. cerevisiae</i>		I G A Q S G L Y I	A D A R D K F V D L Y N
<i>D. melanogaster</i>		- - - S Y L Y M	D V T E L L F - - -
<i>H. sapiens</i>		N E I Q S N I V Y T	G L P P A L F - - -
	<i>HbYX</i>	<i>HbYX</i>	<i>HbF/HbYX</i>

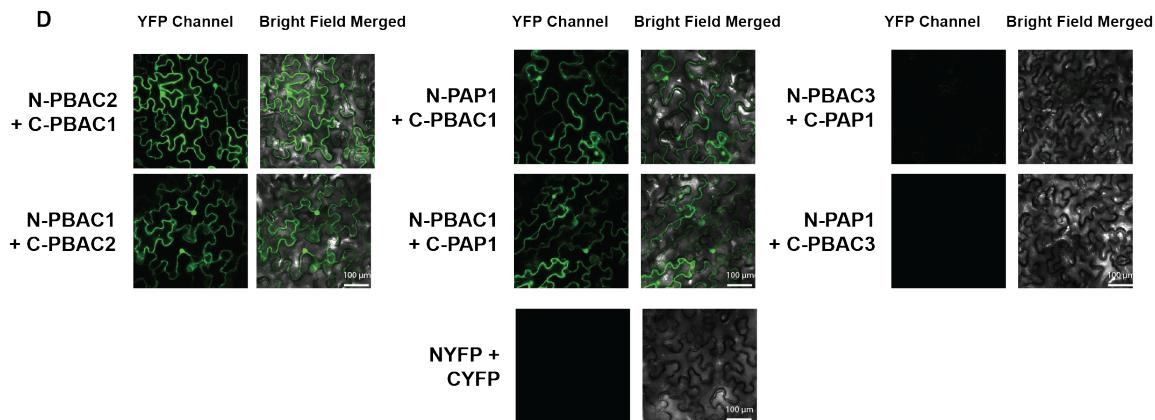
B



C



D



assembly chaperone and that it may interact with one or more of the other putative assembly chaperones. The Y2H showed pairwise interaction for PBAC1 and PBAC2, and pairwise interaction for PBAC3 and PBAC4 regardless of whether they were attached to the GAL4 activating or binding domains (Figure 1.5 C). These interactions were consistent with how the CP assembly orthologs in other organisms interact (Murata *et al.*, 2009). Taken together these data suggested these proteins might form heterodimeric pairs, like other CP assembly chaperone orthologs. Importantly, we saw no interaction from PBAC1 or PBAC2 to PBAC3 or PBAC4 pairs in this

Figure 1.5 (preceding page): PAP1 interacts with the proteasome and with putative assembly chaperone PBAC1. (A) C-terminal sections of protein sequence alignments for PAP1, PBAC1 and PBAC2, show conserved HbYX / HbY/F motifs underlined in red (full alignments in Figures S9, S5, and S6) (B) An anti-HA immunopurification of HA-PAP1 expressed in the PAG1-FLAG background shows affinity enrichment for both PAP1 and the FLAG-tagged PAG1 α -subunit of the CP. A very low level of PBA1 is also observed in the elution. (C) A yeast-two-hybrid analysis of putative plant assembly chaperones shows interaction between PBAC1 and PBAC2 independently of whether PBAC1 is used as bait or prey. Similarly, interaction is also detected between PBAC3 and PBAC4 independently of which protein is used as bait. Interaction between PAP1 PBAC1 is also observed. (D) Bi-molecular Fluorescent Complementation (BiFC) with indicated combinations of PAP1, PBAC1, and PBAC2 fused at their N-terminus with either N or C terminal halves of YFP (N-YFP, C-YFP). BiFC results confirm the results obtained via Y2H in (B) and show that these interactions also occur *in planta* between PBAC1 and PBAC2, and between PBAC1 and PAP1 in either orientation. A control is shown in this figure documenting a lack of auto activation for N-YFP and C-YFP. Additional controls combining N-YFP and C-YFP with their respective N and C terminal PBAC1, PBAC2, PBAC3, and PAP1 controls are shown in Figure S8 along with the full panel from (D) for comparison.

Y2H, suggesting these putative assembly chaperones form separate complexes like their orthologs in mammals and yeast. Interestingly, we observed PAP1 interacting with PBAC1, which suggested that it may be part of the plant assembly chaperone PBAC1/2 complex (Figure 1.5 C).

Because these interactions were analyzed in yeast, which contain their own assembly chaperones that might interfere with interactions between plant subunits, we wanted to verify that PAP1 interacted with PBAC1 *in planta*. To do this we tested several pairwise interactions by bimolecular fluorescent complementation (BiFC) via transient expression in *Nicotiana benthamiana* leaves. Figure 1.5 D shows confocal images of transiently expressed PAP1, PBAC1, PBAC2, and PBAC3 with either the N- or C-terminal half of YFP (NYFP and CYFP respectively) attached as a translational fusion to the N-terminus of each putative assembly chaperone. As controls, PAP1 and PBAC1-3 were tested for auto-activation against NYFP or CYFP alone with no strong auto-activation being observed (Figure S10B). We again observed interactions between PBAC1 and PBAC2, regardless of their attachment to the N or C terminal halves of YFP based on their strong fluorescent intensity (Figure 1.5 D). Additionally, we saw an interaction between PAP1 and PBAC1, showing that PAP1 can interact with PBAC1 *in planta*. Consistent with our Y2H data, there was no evidence of strong interaction between PAP1 and PBAC3 (Figure reffig:papschar

D).

1.5.7. Inhibition of the Proteasome Results in Formation of Distinct Proteasome Sub-Species Containing Putative CP Assembly Chaperones

Upon inhibition with the proteasomal inhibitor MG132, most proteasome genes are coordinately upregulated by a pair of NAC transcription factors, NAC53 and NAC78, as the plants attempt to synthesize more proteasomes in an effort to increase degradation capacity (Gladman *et al.*, 2016). Additionally, after prolonged inhibition the complex becomes heavily ubiquitylated and is turned over via autophagy, likely as a clearing mechanism for inhibited complexes (Marshall *et al.*, 2015). We reasoned that treatment with MG132 might lead to an increase in the population of newly assembling proteasomes, potentially allowing us to observe proteasome assembly intermediates. As such, proteasomes were purified in the presence of ATP from *PAG1::PAG1-FLAG pag1-1* tissue treated for 16 hours with or without MG132. A silver stain of SDS-PAGE-separated proteins affinity-purified from MG132-treated tissue showed successful purification of the 26S proteasome under these conditions (Figure 1.6 A) with the only notable difference in the treated sample being an increased association of PA200, which was also observed by Book *et al.* (Book *et al.*, 2010). An anti-ubiquitin immunoblot of the same samples (Figure 1.6 B), showed

extensive ubiquitylation of the particle consistent with data observed by Marshall et al (Marshall *et al.*, 2015).

Non-denaturing native-PAGE of affinity-purified proteasomes were performed to determine if any novel complexes could be observed. Intriguingly, silver staining of samples affinity-purified from MG132-treated tissue revealed a smear below the canonical CP band (Figure 1.6 B). Additionally, extra bands were observed above the canonical CP band, which may represent stalled intermediates with the alternate capping particle PA200 (Li *et al.*, 2007). To identify the proteins contained in the bands shown in Figure 1.6B, we cut the corresponding gel slices, trypsin-digested the proteins, and then analyzed the resulting peptides via MS/MS. A LFQ-MS approach was used to quantify the proteins present in each sample (Gemperline *et al.*, 2016). Proteins identified in gel slices obtained from both MG132-treated and untreated control samples were clustered based on relative abundance using hierarchical clustering. The clusters were then visualized via a heat map showing that the RP and CP proteins cluster away from one another in these samples (Figure 1.6 C), consistent with their separation by native-PAGE. Furthermore, an outgroup that includes the putative assembly chaperones PBAC1-4, Ump1, and PAP1, plus the accessory factors PA200, PTRE1, as well as ubiquitin, is found in the CP cluster (Figure 1.6 C).

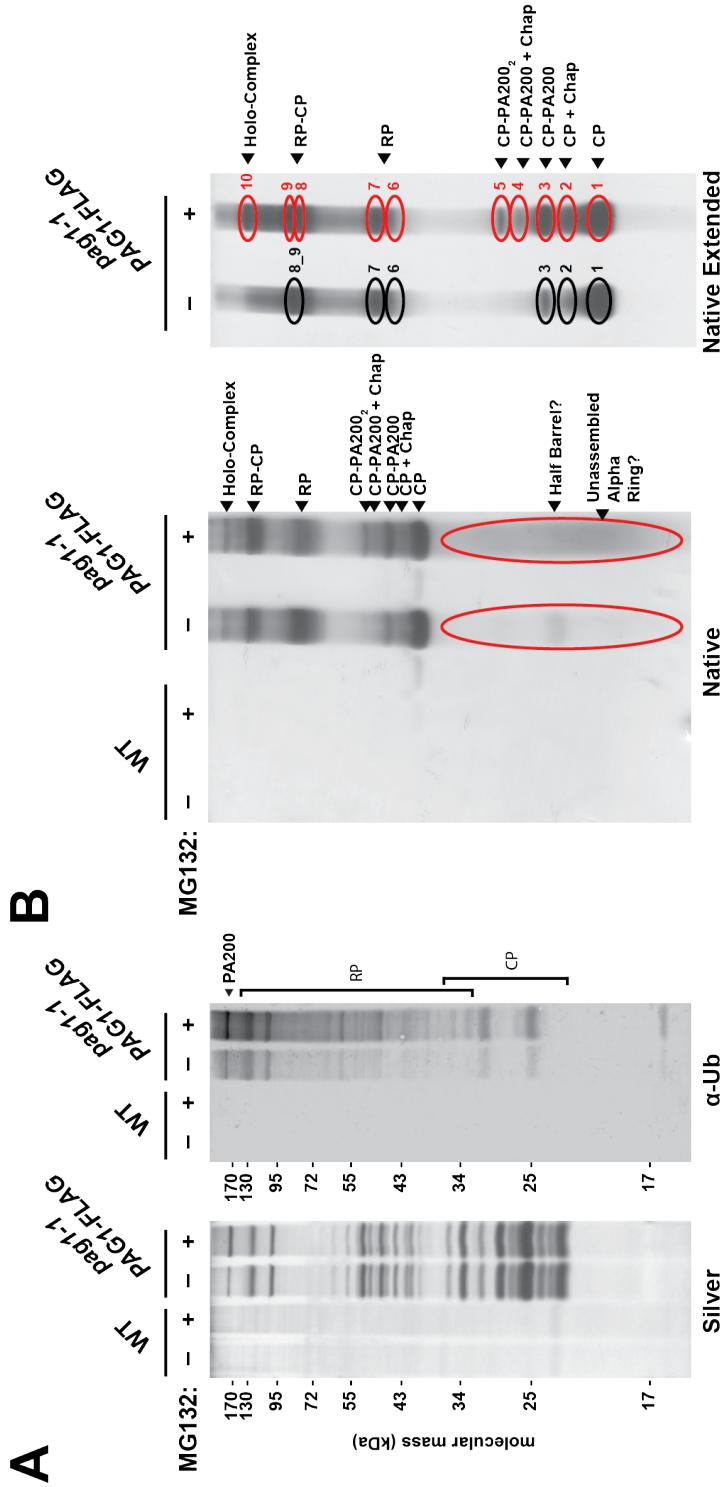


Figure 1.6: Protein composition of PAG1-FLAG affinity-purified proteasomes from tissues treated with and without MG132. (A) A silver stain of proteasomes purified from *Arabidopsis* tissue treated for 24 hours with 50 μ M MG132 (left panel). An increase in PA200 is observed in these samples as judged by silver stain. An anti-ubiquitin immunoblot shows increased ubiquitylation of the complex on the right panel. (B) Native-PAGE analysis of 26S proteasomes affinity purified from MG132-treated (+) and untreated control (-) *Arabidopsis* tissues (left panel). A different banding pattern is observed between MG132-treated and untreated controls, with a smear below, and two bands (4 and 5) above the canonical CP band. An extended gel is shown on the right to better visualize and cut out the indicated bands for mass spectrometry. Cut bands indicated in red were recovered. The proteins contained within these bands were trypsin-digested and subjected to MS/MS analyses.

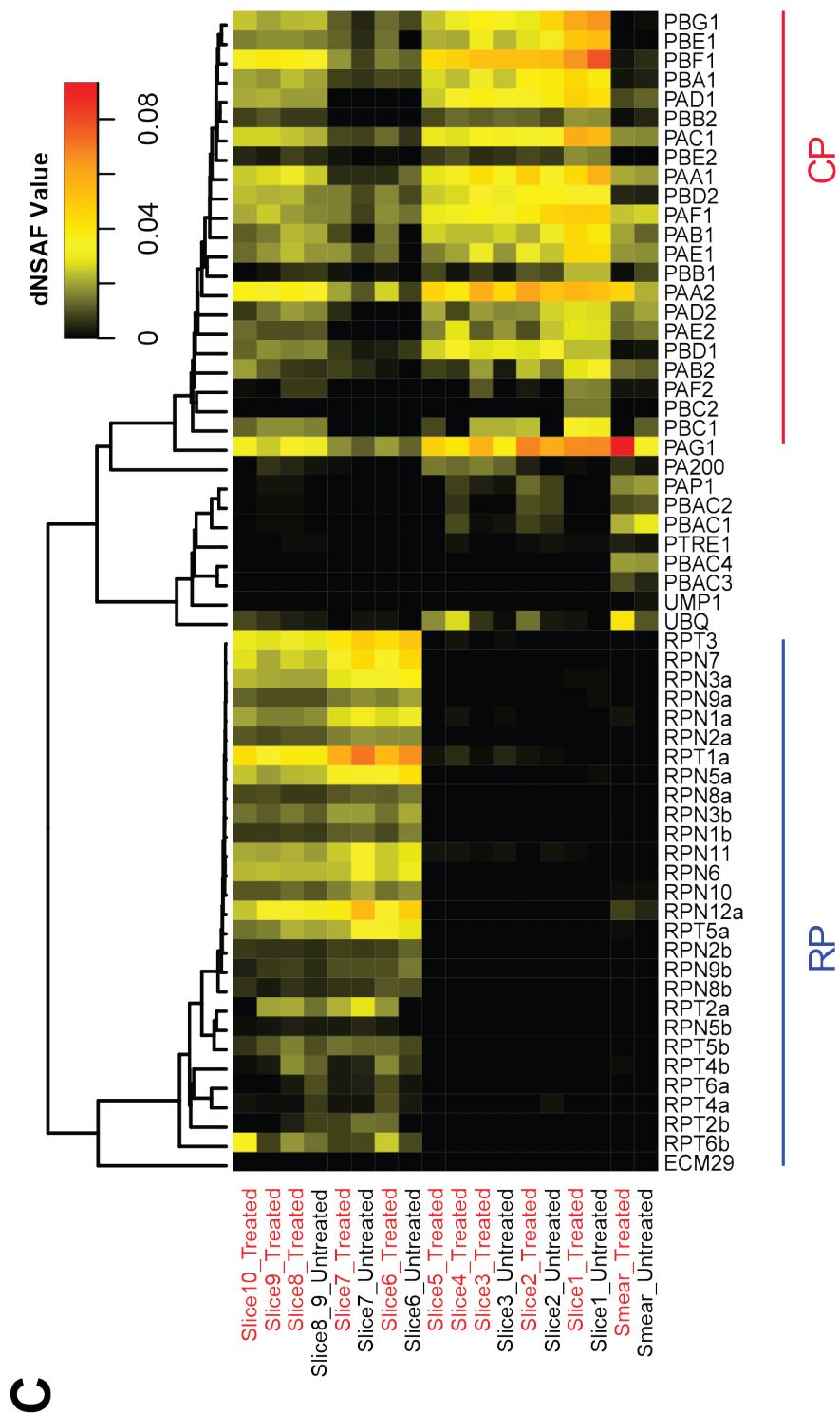


Figure 1.6: Protein composition of PAG1-FLAG affinity-purified proteasomes from tissues treated with and without MG132. (C) A heat map showing LFQ MS/MS analysis (using dNSAF values calculated with Morpheus Spectral Counter (Gemperline *et al.*, 2016)) of the gel slices shown in (B), which were hierarchically clustered using a Pearson's correlation, and ordered by gel slice. Clear grouping of the CP and RP is observed. Fractions containing the unassembled α -rings and putative assembly chaperones are seen from the native gel smears shown in (B). Fractions two and four contain the putative assembly chaperones PBAC1, PBAC2, and PAP1 as well as CP subunits and likely represent pre-assembly intermediates of the CP. PBAC3 and PBAC4 are only seen in the smear.

The proteins identified in gel slices below the canonical CP band in both MG132-treated and untreated control samples (Figure 1.6 B) contained a large abundance of CP α -subunits, with very few β -subunits (Figure 1.6 C). In addition to this set of α -subunits, the previously mentioned CP-specific putative assembly chaperones PBAC1-4 were also observed, as was the plant-specific PAP1 protein and UMP1b. PBAC3 and PBAC4 were only identified in the slice below the canonical CP band. Slices 1-5 contained mostly CP subunits with very few, if any, RP subunits (Figure 1.6 C). Slice 2, above the canonical CP band, from both treated and untreated samples also contained PBAC1, PBAC2, and PAP1, suggesting that this band may represent an assembly intermediate of the plant CP. All CP subunits and the alternate capping particle PA200 were identified in slice 3, suggesting that this band represents CP capped with PA200. Slices 4 and 5 were observed only in proteasomes purified from MG132 treated tissue (Figure 1.6 B, and Table S8), and interestingly, slice 4 contained both the PA200 alternate capping particle, and the assembly chaperones PBAC1, PBAC2, and PAP1 along with most of the CP subunits, suggesting that this may be a PA200 containing assembly intermediate. Slice 5 likely represented a doubly capped PA200-CP-PA200 species, as no other putative assembly chaperones were present, and a strong signal for PA200, and the various CP subunits was observed.

In contrast to slices 1-5, most RP subunits were identified in slices 6-10. Based on size, slices 6-7 likely represented free RP complexes that have some associating α -subunits, while slices 8-9 likely represent doubly or singly capped RP-CP complexes (Book *et al.*, 2010). It was difficult to discriminate between singly and doubly capped proteasomes in these native gel assays, and we did not observe discrete bands for these complexes. Slice 10 likely represented the holo-complex of the 26S proteasome (Book *et al.*, 2010), and notably is the only slice in which the CP-RP stabilizing protein ECM29 was identified; however, it was detected at low abundance. Very low levels of the putative assembly chaperones PBAC1, PBAC2, and PAP1 were detected in these higher slices. Increased amounts of ubiquitin were detected in slices 1-5 for proteasomes that were affinity-purified from MG132-treated tissue, which also contained the putative assembly chaperones PBAC1, PBAC2, and PAP1. No ubiquitin receptors were identified in these slices, as expected, which suggests that ubiquitin is covalently attached; however we were unable to identify any canonical ubiquitylation footprints (GlyGly) (Wang *et al.*, 2006) on peptides corresponding to any proteasome subunits in these samples. Taken together, these data suggest that the novel upper novel slices and lower smears observed might represent stalled assembly intermediates, and that they may be actively ubiquitylated and turned over to prevent the formation of aberrant complexes.

1.6. Discussion

Here we have shown that we can affinity purify proteasomes from plants expressing FLAG-tagged RP subunits RPT4a and RPT4b, and that these preparations are of sufficient purity to enable analysis by MS/MS. Using a LFQ-MS/MS strategy, we were able to show that these RP-based affinity purifications identify the core components of the 26S proteasome, and are comparable to our previously developed CP-based affinity purification. Interestingly the only major difference between the RPT4a and RPT4b purifications was the statistical enrichment for each respective isoform when purifications were performed in the absence of ATP. Even with this enrichment, no other isoforms or associated proteins were preferentially enriched. For RPT4a/b purifications performed in the presence of ATP, enrichment for each respective subunit was observed; however, this difference was not statistically significant. These data imply that plants do not assemble particular proteasome subunit isoforms into specific sub-particles or isotypes at the young-seedling stage. While most plant proteasome subunits are encoded as paralogs, mammals instead have only a few subunit paralogs. $\beta 1$, $\beta 2$, and $\beta 5$ are replaced in immune tissue with $\beta 1i$, $\beta 2i$, and $\beta 5i$ to form the immunoproteasome isotype (Nandi *et al.*, 1996), and in thymus tissue the immunoproteasome contains $\beta 5t$ instead of $\beta 5i$ forming the thymoproteasome isotype (Murata *et al.*, 2007). Outside of these catalytic subunits,

the only other isoform switch involves replacing the $\alpha 4$ subunit with a variant only expressed in testes forming a testes specific proteasome isotype (Belote *et al.*, 1998; Uechi *et al.*, 2014). While there is no evidence for *Arabidopsis* assembling proteasomes into isoform specific sub-types, duplication of the RPT family has also occurred in monocots, with rice having duplicated subunits for RPT1, 2, 4 and 5. This suggests that there may be some evolutionary benefit to having multiple copies of the RPT family (Shibahara *et al.*, 2004). Interestingly, proteomic analysis of bran and callus rice tissue showed some small differences in isoform expression suggesting that analysis of different *Arabidopsis* developmental stages may provide additional insight into potential isoform-specific functions (Shibahara *et al.*, 2004). In conclusion, we did not find proteasome isotypes at the young seedling stage that are specific to either RPT4 isoform. In other organisms, proteasome sub-types seem to be found in specific tissues or cell-types (Belote *et al.*, 1998; Murata *et al.*, 2007; Nandi *et al.*, 1996; Uechi *et al.*, 2014). Therefore, affinity purifications of proteasomes from different plant tissues and developmental stages may be an intriguing next step to explore the possible existence of plant proteasome isotypes.

In conjunction with our previously developed CP affinity purification (Book *et al.*, 2010), this newly developed RP affinity purification allowed us to specifically enrich for either the CP or RP sub-complexes when purifications are performed in

the absence of ATP. LFQ-MS/MS and statistical analyses of these affinity-purified proteasome sub-complexes enabled identification of sub-complex-specific PAPs. The most predominant RP-specific PAPs present in both RPT4a and RPT4b affinity-purified samples are likely orthologs of the RP assembly chaperones NAS2, NAS6, and HSM3. We were unable to identify any proteins that were likely orthologs for the RP assembly chaperone RPN14. In yeast, RPN14 and NAS6 bind together in a complex, and only double mutants show a strong growth defect in growth, suggesting that together they perform redundant functions (Funakoshi *et al.*, 2009). These data suggest that that plant proteasome assembly has diverged from the yeast and mammalian pathways and may no longer require RPN14 to assemble the RP (Funakoshi *et al.*, 2009). There are no obvious orthologs for RPN14 in the *Arabidopsis* genome (Book *et al.*, 2010) which supports this hypothesis. The exact role of NAS2, NAS6, and HSM3 in plant proteasome assembly remains to be determined; however, their specific association with the RP suggests that they may perform similar functions in RP assembly as in mammals and yeast. It will be intriguing to test if these RP assembly orthologs form similar modules with specific RP subunits like their yeast and mammalian counterparts (Nas2 with Rpt4 and 5, Nas6 with Rpt3, Hsm3 with Rpt1,2 and Rpn1, and Rpn14 with Rpt6) (Park *et al.*, 2010).

Several PAPs identified here were enriched in affinity purifications targeting

both the RP and CP. Unsurprisingly, the alternate capping particle PA200 was enriched in our CP-specific purifications. This protein is known to interact with α -subunits of the CP (Ortega *et al.*, 2005). However, we detect PA200 in our RP affinity purifications as well, which suggests that plants form a hybrid RP-CP-PA200 complex. This hybrid complex has been observed in yeast, and is speculated to aide in release of degraded substrate peptides (Ortega *et al.*, 2005; Schmidt *et al.*, 2005; Ustrell *et al.*, 2002). In plants however, PA200 mutants display a remarkably wild-type phenotype (Book *et al.*, 2010), which suggests that this hybrid complex is not essential in *Arabidopsis*. ECM29 was also found to be enriched as compared to mock controls in both CP and RP targeted purifications. ECM29 is known in yeast to bind the CP subunit PRE10 (named PAG1 in *Arabidopsis*), and the RP subunit Rpt5 (Wani *et al.*, 2016). In yeast, ECM29 is one of the only PAPs known to bind both the RP and the CP (Lehmann *et al.*, 2010), consistent with our observations. Ecm29 has been shown to inhibit the yeast 26S proteasome, and may be a checkpoint for damaged or aberrantly assembled complexes (Lehmann *et al.*, 2010). While ECM29 likely performs a similar role in *Arabidopsis*, its function in plants remains to be investigated.

A recently characterized plant ortholog of proteasome inhibitor 31 (PI31), named Proteasome Regulator 1 or PTRE1 (Yang *et al.*, 2016), was identified as a CP-specific

PAP, providing confirmation of PTRE1s association with the CP of the 26S proteasome. However, the most predominant CP-specific PAPs identified here were putative orthologs of the CP assembly chaperones PBAC1-4, UMP1, and the plant specific protein PAP1. In our analyses, searches for PBAC1-4 in the *Arabidopsis* genome using assembly chaperones from yeast (Pba1-4) or humans (PAC1-4) failed with BLAST alone. This is not surprising given that orthologs from different kingdoms typically shared <25% identity, with fungi and animal forms of these chaperones typically sharing <15% sequence identity (Table S3-S6) despite having similar functions. However two independent groups identified orthologs to the assembly chaperones PBAC1 and PBAC2 using an iterative PSI-BLAST based approach with the human orthologs PAC1 and PAC2 (Kusmierczyk *et al.*, 2011; Le Tallec *et al.*, 2007). PSI-BLAST has several advantages over BLASTp in that it is typically more sensitive, and can be useful when identifying protein like PBAC1-4 that likely share structural homology but share little sequence homology (Altschul *et al.*, 1997; Aravind and Koonin, 1999; Kusmierczyk *et al.*, 2011; Yashiroda *et al.*, 2008). We were able to identify PBAC1 and PBAC2 as the top *Arabidopsis* PSI-BLAST hits for human PAC1 and PAC2 (Table S3B), respectively, repeating the finding by the both Le Tallec *et al.* (2007), Kusmierczyk *et al.* (2011). Additionally, we were able to extend this analysis, with both PBAC3 and PBAC4 being top PSI-BLAST hits for

human PAC3 and PAC4 (Table S3B). PSI-BLAST analysis with the yeast assembly chaperones Pba1-4 failed to identify the putative *Arabidopsis* orthologs PBAC1-4, suggesting that the plant CP assembly system may be different from the yeast system. Sequence-based InterProScan analyses showed that PBAC2-4 contained PAC2-4 domains (Table S3A), further supporting these assignments. InterProScan analyses showed that PBAC1 contained a PAC2 domain; however, the E-value for PBAC1 containing this domain is much lower than the assignment of this domain to PBAC2 (3.4E-6 vs 1.0E-29), and is much lower than the other assigned domains. Despite this disagreement, PBAC1 was the top hit in the *Arabidopsis* genome when searching via PSI-BLAST with human PAC1, as performed by ourselves and others, and for these reasons we have decided to use the name PBAC1 for this identified protein (Kusmierczyk *et al.*, 2011; Le Tallec *et al.*, 2007). Sequence alignments of PBAC1 and PBAC2 also showed conserved HbYX and HbY/F motifs respectively that are also found in their human counterparts. Sequence analysis alone, however, is insufficient to conclude that these are assembly chaperones. There is strong divergence in sequence identity between the plant, yeast, and animal sequences (alignments and identity matrices in Figures S5–S8); therefore, additional characterizations will be necessary to determine what role if any they have in plant CP assembly. Despite this divergence, these bioinformatics analyses, along with

prior literature (Kusmierczyk *et al.*, 2011; Le Tallec *et al.*, 2007), help support our identification of these CP-specific PAPs as PBAC1-4.

Phylogenetic analysis of the putative plant assembly chaperones orthologs PBAC1-4, and plant specific protein PAP1, showed that PBAC2-4 formed clades with their respective animal counterparts (Figures S11A and B); however, PBAC2 was the only protein to form a clear clade with all sequences from animals, yeast, and plants. PBAC1, 3, and 4 from yeast did not form clades with either the plant or animal sequences, suggesting that they are more different in sequence as compared to the plant and animal versions. In this analysis PBAC1 from yeast, mammals and plants was a strong exception in that it did not form a strong clade with any of their respective orthologs (Figures S11A). These data may not be surprising given the low sequence homology shared between yeast and animal PBAC1, with sequence identity at <13%. This low % identity makes phylogenetic analysis of these putative PBAC1 orthologs particularly challenging, therefore we cannot draw strong conclusions about PBAC1 orthology based on this analysis alone.

In yeast and mammals, Pba1/PAC1 and Pba2/PAC2 form a heterodimeric pair. Separately, Pba3/PAC3 and Pba4/PAC4 also form a heterodimeric pair. If plant PBAC1-4 are acting similarly, we would expect that PBAC1 should interact with PBAC2 and that PBAC3 should interact with PBAC4. Indeed, our Y2H analyses

shows these interactions (Figure 1.5 C) suggesting that they each form a heterodimer like their respective orthologs in yeast and mammals (Kunjappu and Hochstrasser, 2014). Additionally, in our Y2H we failed to detect interactions between these two putative heterodimeric pairs as both PBAC1 and PBAC2 failed to interact with PBAC3 and PBAC4, which suggests that these proteins form two separate complexes. These data were also re-capitulated in our BiFC analyses (Figure 1.5 D) suggesting that these interactions occur *in planta*; however, the interaction between PBAC3 and PBAC4 was not tested via BiFC.

We were able to confirm the interaction of the CP-specific PAP1 protein with the proteasome by IP with an HA tagged version, identifying the CP subunit PAG1. The presence of a conserved HbYX motif in PAP1 (Figure 1.5 A), which was also found in the putative assembly chaperone PBAC1, suggested that PAP1 might play a role in plant proteasome assembly. In yeast, the Pba1 and Pba2 proteins both have HbYX motifs important for interacting with the CP α -ring; however orthologs of Pba2 in higher eukaryotes lack HbYX motifs, and instead have a HbY/F motif that lacks the last variable C-terminal amino acid (Kusmierczyk *et al.*, 2011). Consistent with this, PBAC1 has a HbYX motif conserved in plants, and PBAC2 has a HbY/F motif conserved in all plants analyzed, except *Physcomitrella patens* (which has a HbFX motif, Figure 1.5 A), suggesting the plant-CP-assembly pathway is more

similar to that of higher eukaryotes than that of fungi. If PAP1 is functioning as a CP assembly chaperone it might be acting in concert with PBAC1, given their interaction (Figure 1.5 C and D). In one scenario, PAP1 could replace PBAC2 in the PBAC1/2 heterodimer, or in an alternative scenario, PAP1 may be forming a novel trimeric complex with both PBAC1 and PBAC2. Despite our evidence that PAP1 contains a conserved HbYX motif, and interacts with the plant assembly chaperone PBAC1 via both Y2H and *in planta* by BiFC, we cannot rule out that PAP1 functions outside of plant proteasome assembly.

While an interaction network of these putative plant proteasome assembly chaperones provides some insight into how these proteins organize, our analysis of complexes affinity purified from PAG1-FLAG tissue treated with the potent and specific proteasome inhibitor MG132 shows the formation of novel sub-complexes that contain these putative chaperones. In general, two classes of species appeared in the sample affinity purified from tissue treated with MG132 as analyzed by native-PAGE. The first is a smear running at a faster pace below the canonical CP species which was increased in intensity upon inhibitor treatment. The second class includes discrete species migrating more slowly than the canonical CP band (Figure 1.6 B).

In the untreated sample, native-PAGE showed a very faint band (as opposed to

a smear) below the canonical CP band (Figure 1.6 B, labeled Half Barrel?). This may represent a proteasome half-barrel known as a 15S precursor complex that also typically runs below the CP band in native-PAGE analyses and contains partially assembled α -subunits along with UMP1 (Kock *et al.*, 2015). Indeed, our MS/MS analyses of this faint band identified UMP1, α -subunits and the putative assembly chaperones PBAC1-4 and PAP1. While mostly α -subunits were present, some β -subunits were still identified at lower levels. In yeast, these 15S complexes lack the β 7 subunit (Kock *et al.*, 2015). We still detected the β 7 subunit PBG1; however, it was present at a 50-fold lower level as compared to the CP band in our untreated samples (Table S5). Native-PAGE analysis of the affinity preparations from MG132-treated tissue showed a smear instead of the faint band seen in the untreated samples. MS/MS analyses of this smear showed that it also contained unassembled α -subunits, UMP1, and the suite of putative assembly chaperones PBAC1-4. We could not observe a discrete band similar to the 15S species from yeast but this may be due to the fact that we are inhibiting the complex and are capturing a wider array of partially assembled products, or that these partially assembled products are less stable in plants as compared to yeast. Indeed, increased levels of ubiquitin were found in this fraction even though no ubiquitin binding proteins were detected (as expected), suggesting that these sub-species may be aberrant and are likely turned

over, either via autophagy or other active proteasomes. PBAC3, PBAC4, and UMP1 are exclusively detected in the untreated and treated smears below the CP band, which is consistent with PBAC3 and PBAC4 preferentially binding free α -subunits at early stages of proteasome assembly, as established in both yeast and mammals (Kunjappu and Hochstrasser, 2014), and with UMP1 being degraded at later stages of proteasome assembly (Ramos *et al.*, 1998).

The new species observed above the canonical CP band upon MG132 treatment contained large amounts of PA200, suggesting that these are likely singly- and doubly-capped versions of the CP (CP-PA200, and PA200-CP-PA200, respectively). The first band above the CP was previously identified as CP-PA200 by our lab (Book *et al.*, 2010); however, this doubly capped complex and additional bands observed in affinity purified complexes from MG132 inhibited tissue have not been seen before in plants. Recently, other researchers have found that human cell lines treated with the proteasome inhibitor bortezomib induced the formation of a species above the canonical CP band, which also contained PA200 (Welk *et al.*, 2016) suggesting that inhibition of the complex increases CP-PA200 binding, and that this process is conserved between animals and plants. The function of increased PA200 binding to the CP is still unclear in both plants and animals. It could be that inhibition of the complex causes PA200 to associate and help release peptides from

stalled proteasomes. Alternatively, given PA200s role in assembly (Savulescu and Glickman, 2011), it could represent increased biogenesis of the CP sub-complex in response to inhibition.

Intriguingly, high levels of PBAC1, PBAC2, and PAP1 were found in the native-PAGE bands directly above the canonical CP band (slice 2 Figure 1.6 B and Figure 1.6 C) in samples affinity purified from both untreated and MG132 treated tissues. This suggests that PBAC1, PBAC2, and PAP1 may be acting at later stages of proteasome assembly, given their association with species that run higher than mature CP. While the assembly chaperones Pba1/PAC1 and Pba2/PAC2 in yeast and mammals typically do not associate with mature forms of the CP, the related archaeal assembly chaperones PbaA and PbaB are found loosely associated with mature CP when its active sites are bound with inhibitors (Kusmierczyk *et al.*, 2011). The data gathered here suggest that PBAC1, PBAC2, and PAP1 can bind mature CP and suggest that plant CP assembly at later stages may be more similar to archaeal proteasome assembly. Taken together these larger-sized bands may represent stalled assembly intermediates of the mature proteasome still carrying their assembly chaperones. Like the previously mentioned smears, increased levels of ubiquitin are found in these fractions even though no ubiquitin binding proteins were detected, suggesting that these sub-species may also be aberrant and are likely turned over. While the

exact role of PAP1 in plant proteasome assembly, if any, still needs to be determined, the data presented here suggest that it may be involved in late-stage proteasome assembly, as it co-migrates with PBAC1 and PBAC2. Figure 1.7 illustrates a possible model of CP assembly that is consistent with the data collected here.

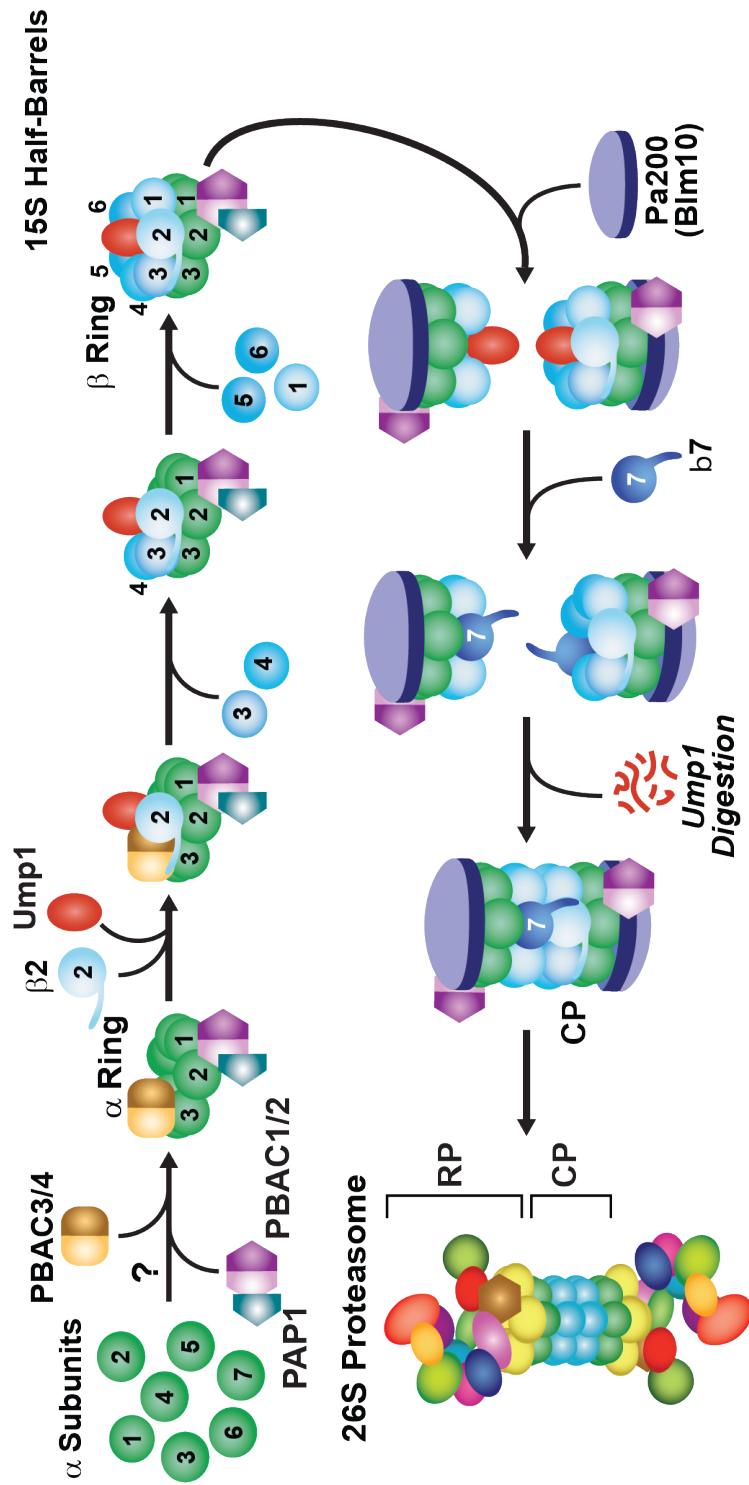


Figure 1.7: Putative Model of 20S Proteasome Assembly in *Arabidopsis*. A model of 20S proteasome assembly adapted from (Murata *et al.*, 2009).

In summary, the data presented herein show that we were able to efficiently purify the 26S proteasome via both the CP and RP. The newly-developed RP-based purification allowed us to identify CP and RP specific PAPs, several of which may be involved in the assembly of their respective sub-complexes. The fact that the putative CP assembly chaperones were only identified in our CP-based affinity purifications, and that the putative RP assembly chaperones were only identified in our RP-based affinity purifications, suggest that only mature RP is purified with our CP affinity preparation, and that only mature CP is purified with our RP affinity purification. Our follow up analysis builds an interaction network of PBAC1-4 and PAP1, and identifies novel sub-species that contain these putative assembly chaperones when analyzing sub-complexes affinity purified from MG132-treated tissue. Taken together, these data suggest that UMP1, PBAC3, and PBAC4 may be performing a role in the early biogenesis of the CP, while PBAC1, PBAC2, and possibly PAP1 are playing a role in later stages of CP assembly. While we have been able to identify the likely players in the plant proteasome assembly pathways for both the RP and CP, additional functional analyses of these proteins will be needed to demonstrate their roles as putative assembly chaperones, and to fully elucidate the plant proteasome assembly pathway.

1.7. Literature Cited

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J.** (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–402.
- Aravind, L. and Koonin, E.V.** (1999). Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. *J. Mol. Biol.* **287**: 1023–40.
- Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J.M., Villa, E., Baumeister, W., and Förster, F.** (2012). Near-atomic resolution structural model of the yeast 26S proteasome. *Proc. Natl. Acad. Sci. U.S.A.* **109**: 14870–14875.
- Belote, J.M., Miller, M., and Smyth, K.A.** (1998). Evolutionary conservation of a testes-specific proteasome subunit gene in *Drosophila*. *Gene* **215**: 93–100.
- Berardini, T.Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., and Huala, E.** (2015). The arabidopsis information resource: Making and mining the gold standard annotated reference plant genome. *Genesis* **53**: 474–485.
- Besche, H.C., Haas, W., Gygi, S.P., and Goldberg, A.L.** (2009). Isolation of mammalian 26S proteasomes and p97/VCP complexes using the ubiquitin-like domain from HHR23B reveals novel proteasome-associated proteins. *Biochemistry (Mosc.)* **48**: 2538–49.
- Book, A.J., Gladman, N.P., Lee, S.S., Scalf, M., Smith, L.M., and Vierstra, R.D.** (2010). Affinity purification of the Arabidopsis 26S proteasome reveals a diverse array of plant proteolytic complexes. *J. Biol. Chem.* **285**: 25554–25569.
- Book, A.J., Smalle, J., Lee, K.H., Yang, P., Walker, J.M., Casper, S., Holmes, J.H., Russo, L.A., Buzzinotti, Z.W., Jenik, P.D., and Vierstra, R.D.** (2009). The RPN5 subunit of the 26S proteasome is essential for gametogenesis, sporophyte development, and complex assembly in Arabidopsis. *Plant Cell* **21**: 460–78.
- Brukhin, V., Gheyselinck, J., Gagliardini, V., Genschik, P., and Grossniklaus, U.** (2005). The RPN1 subunit of the 26S proteasome in Arabidopsis is essential for embryogenesis. *Plant Cell* **17**: 2723–37.
- Chevallet, M., Luche, S., and Rabilloud, T.** (2006). Silver staining of proteins in polyacrylamide gels. *Nat. Protoc.* **1**: 1852–8.

- Clough, S.J. and Bent, A.F.** (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**: 735–43.
- Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M.** (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**: 2513–26.
- Cox, J. and Mann, M.** (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**: 1367–72.
- Dange, T., Smith, D.M., Noy, T., Rommel, P.C., Jurzitz, L., Cordero, R.J., Legende, A., Finley, D., Goldberg, A.L., and Schmidt, M.** (2011). The Blm10 protein promotes proteasomal substrate turnover by an active gating mechanism. *J. Biol. Chem.* **286**: 42830–42839.
- Edelheit, O., Hanukoglu, A., and Hanukoglu, I.** (2009). Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* **9**: 61.
- Elsasser, S., Chandler-Militello, D., Muller, B., Hanna, J., and Finley, D.** (2004). Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome. *J. Biol. Chem.* **279**: 26817–22.
- Elsasser, S., Gali, R.R., Schwickart, M., Larsen, C.N., Leggett, D.S., Muller, B., Feng, M.T., Tubing, F., Dittmar, G.A., and Finley, D.** (2002). Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* **4**: 725–30.
- Farmer, L.M., Book, A.J., Lee, K.H., Lin, Y.L., Fu, H., and Vierstra, R.D.** (2010). The RAD23 family provides an essential connection between the 26S proteasome and ubiquitylated proteins in *Arabidopsis*. *Plant Cell* **22**: 124–42.
- Fatimababy, A.S., Lin, Y.L., Usharani, R., Radjacommare, R., Wang, H.T., Tsai, H.L., Lee, Y., and Fu, H.** (2010). Cross-species divergence of the major recognition pathways of ubiquitylated substrates for ubiquitin/26S proteasome-mediated proteolysis. *FEBS J.* **277**: 796–816.
- Finley, D.** (2009). Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **78**: 477–513.

- Fu, H., Doelling, J.H., Arendt, C.S., Hochstrasser, M., and Vierstra, R.D.** (1998). Molecular organization of the 20S proteasome gene family from *Arabidopsis thaliana*. *Genetics* **149**: 677–92.
- Funakoshi, M., Tomko, R.J., J., Kobayashi, H., and Hochstrasser, M.** (2009). Multiple assembly chaperones govern biogenesis of the proteasome regulatory particle base. *Cell* **137**: 887–99.
- Gamborg, O.L., Miller, R.A., and Ojima, K.** (1968). Nutrient requirements of suspension cultures of soybean root cells. *Exp. Cell. Res.* **50**: 151–8.
- Gelvin, S.B.** (2003). Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev.* **67**: 16–37, table of contents.
- Gemperline, D.C., Scalf, M., Smith, L.M., and Vierstra, R.D.** (2016). Morpheus Spectral Counter: A computational tool for label-free quantitative mass spectrometry using the Morpheus search engine. *Proteomics* **16**: 920–4.
- Gladman, N.P., Marshall, R.S., Lee, K.H., and Vierstra, R.D.** (2016). The Proteasome Stress Regulon Is Controlled by a Pair of NAC Transcription Factors in *Arabidopsis*. *Plant Cell* **28**: 1279–96.
- Gonnet, G.H., Cohen, M.A., and Benner, S.A.** (1992). Exhaustive matching of the entire protein sequence database. *Science* **256**: 1443–5.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S.** (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**: D1178–86.
- Hirano, Y., Kaneko, T., Okamoto, K., Bai, M., Yashiroda, H., Furuyama, K., Kato, K., Tanaka, K., and Murata, S.** (2008). Dissecting beta-ring assembly pathway of the mammalian 20S proteasome. *EMBO J.* **27**: 2204–13.
- Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesceat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., and Hunter, S.** (2014). InterProScan 5 genome-scale protein function classification. *Bioinformatics* **30**: 1236–40.
- Julio, S. and Jose, S.S.J.** (2006). *Arabidopsis Protocols*, Second Edition, volume 323. Humana Press.

- Kock, M., Nunes, M.M., Hemann, M., Kube, S., Dohmen, R.J., Herzog, F., Ramos, P.C., and Wendler, P.** (2015). Proteasome assembly from 15S precursors involves major conformational changes and recycling of the Pba1-Pba2 chaperone. *Nat. Commun.* **6**: 6123.
- Kunjappu, M.J. and Hochstrasser, M.** (2014). Assembly of the 20S proteasome. *Biochim. Biophys. Acta* **1843**: 2–12.
- Kusmierczyk, A.R., Kunjappu, M.J., Kim, R.Y., and Hochstrasser, M.** (2011). A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding. *Nat. Struct. Mol. Biol.* **18**: 622–9.
- Laemmli, U.K.** (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227**: 680–685.
- Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., and Martin, A.** (2012). Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**: 186–191.
- Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W.** (2012). Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. U.S.A.* **109**: 1380–1387.
- Le Tallec, B., Barrault, M.B., Courbeyrette, R., Guérois, R., Marsolier-Kergoat, M.C., and Peyroche, A.** (2007). 20S proteasome assembly is orchestrated by two distinct pairs of chaperones in yeast and mammals. *Mol. Cell* **27**: 660–674.
- Leggett, D.S., Glickman, M.H., and Finley, D.** (2005). Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast. *Methods Mol. Biol.* **301**: 57–70.
- Leggett, D.S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R.T., Walz, T., Ploegh, H.L., and Finley, D.** (2002). Multiple associated proteins regulate proteasome structure and function. *Mol. Cell* **10**: 495–507.
- Lehmann, A., Niewienda, A., Jechow, K., Janek, K., and Enenkel, C.** (2010). Ecm29 fulfills quality control functions in proteasome assembly. *Mol. Cell* **38**: 879–88.
- Li, X., Kusmierczyk, A.R., Wong, P., Emili, A., and Hochstrasser, M.** (2007). beta-Subunit appendages promote 20S proteasome assembly by overcoming an Ump1-dependent checkpoint. *EMBO J.* **26**: 2339–49.

- Lin, Y.L., Sung, S.C., Tsai, H.L., Yu, T.T., Radjacompare, R., Usharani, R., Fatimababy, A.S., Lin, H.Y., Wang, Y.Y., and Fu, H.** (2011). The defective proteasome, not substrate recognition function, is responsible for the null phenotypes of the Arabidopsis proteasome subunit RPN10. *Plant Cell* **23**: 2754–2773.
- Liu, C.W., Li, X., Thompson, D., Wooding, K., Chang, T.L., Tang, Z., Yu, H., Thomas, P.J., and DeMartino, G.N.** (2006). ATP binding and ATP hydrolysis play distinct roles in the function of 26S proteasome. *Mol. Cell* **24**: 39–50.
- Livneh, I., Cohen-Kaplan, V., Cohen-Rosenzweig, C., Avni, N., and Ciechanover, A.** (2016). The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell Res.* **26**: 869–85.
- Marques, A.J., Glanemann, C., Ramos, P.C., and Dohmen, R.J.** (2007). The C-terminal extension of the beta7 subunit and activator complexes stabilize nascent 20 S proteasomes and promote their maturation. *J. Biol. Chem.* **282**: 34869–76.
- Marshall, R.S., Gemperline, D.C., and Vierstra, R.D.** (2017). Isolation of Plant Organelles and Structures, volume 1 of *Methods in Molecular Biology*. Humana Press, Methods in Molecular Biology, 1 edition.
- Marshall, R.S., Li, F., Gemperline, D.C., Book, A.J., and Vierstra, R.D.** (2015). Autophagic Degradation of the 26S Proteasome Is Mediated by the Dual ATG8/Ubiqutin Receptor RPN10 in Arabidopsis. *Mol. Cell* **58**: 1053–66.
- Martin, K., Kopperud, K., Chakrabarty, R., Banerjee, R., Brooks, R., and Goodin, M.M.** (2009). Transient expression in Nicotiana benthamiana fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta. *Plant J.* **59**: 150–62.
- Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabkin, S., Guo, N., Muruganujan, A., Doremieux, O., Campbell, M.J., Kitano, H., and Thomas, P.D.** (2005). The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res.* **33**: D284–8.
- Murashige, T. and Skoog, F.** (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiol. Plant.* **15**: 473–497.
- Murata, S., Sasaki, K., Kishimoto, T., Niwa, S., Hayashi, H., Takahama, Y., and Tanaka, K.** (2007). Regulation of CD8+ T cell development by thymus-specific proteasomes. *Science* **316**: 1349–53.

- Murata, S., Yashiroda, H., and Tanaka, K.** (2009). Molecular mechanisms of proteasome assembly. *Nat. Rev. Mol. Cell Biol.* **10**: 104–15.
- Nandi, D., Jiang, H., and Monaco, J.J.** (1996). Identification of MECL-1 (LMP-10) as the third IFN-gamma-inducible proteasome subunit. *J. Immunol.* **156**: 2361–4.
- Ortega, J., Heymann, J.B., Kajava, A.V., Ustell, V., Rechsteiner, M., and Steven, A.C.** (2005). The axial channel of the 20S proteasome opens upon binding of the PA200 activator. *J. Mol. Biol.* **346**: 1221–7.
- Paraskevopoulos, K., Kriegenburg, F., Tatham, M.H., Rösner, H.I., Medina, B., Larsen, I.B., Brandstrup, R., Hardwick, K.G., Hay, R.T., Kragelund, B.B., Hartmann-Petersen, R., and Gordon, C.** (2014). Dss1 is a 26S proteasome ubiquitin receptor. *Mol. Cell* **56**: 453–461.
- Park, S., Tian, G., Roelofs, J., and Finley, D.** (2010). Assembly manual for the proteasome regulatory particle: the first draft. *Biochem. Soc. Trans.* **38**: 6–13.
- Pathare, G.R., Nagy, I., Bohn, S., Unverdorben, P., Hubert, A., Korner, R., Nickell, S., Lasker, K., Sali, A., Tamura, T., Nishioka, T., Forster, F., Baumeister, W., and Bracher, A.** (2012). The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. *Proc. Natl. Acad. Sci. U.S.A.* **109**: 149–54.
- Ramos, P.C., Hockendorff, J., Johnson, E.S., Varshavsky, A., and Dohmen, R.J.** (1998). Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* **92**: 489–99.
- Rappsilber, J., Ishihama, Y., and Mann, M.** (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**: 663–70.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P.** (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**: 539–42.
- Russell, J.D., Scalf, M., Book, A.J., Ladrón, D.T., Vierstra, R.D., Smith, L.M., and Coon, J.J.** (2013). Characterization and quantification of intact 26S proteasome proteins by real-time measurement of intrinsic fluorescence prior to top-down mass spectrometry. *PLoS One* **8**: e58157.

- Sadre-Bazzaz, K., Whitby, F.G., Robinson, H., Formosa, T., and Hill, C.P.** (2010). Structure of a Blm10 complex reveals common mechanisms for proteasome binding and gate opening. *Mol. Cell* **37**: 728–735.
- Saveliev, S.V., Woodroffe, C.C., Sabat, G., Adams, C.M., Klaubert, D., Wood, K., and Urh, M.** (2013). Mass spectrometry compatible surfactant for optimized in-gel protein digestion. *Anal. Chem.* **85**: 907–14.
- Savulescu, A.F. and Glickman, M.H.** (2011). Proteasome activator 200: the heat is on. *Mol. Cell. Proteomics* **10**: R110 006890.
- Schmidt, M., Haas, W., Crosas, B., Santamaria, P.G., Gygi, S.P., Walz, T., and Finley, D.** (2005). The HEAT repeat protein Blm10 regulates the yeast proteasome by capping the core particle. *Nat. Struct. Mol. Biol.* **12**: 294–303.
- Schreiner, P., Chen, X., Husnjak, K., Randles, L., Zhang, N., Elsasser, S., Finley, D., Dikic, I., Walters, K.J., and Groll, M.** (2008). Ubiquitin docking at the proteasome through a novel pleckstrin-homology domain interaction. *Nature* **453**: 548–52.
- Shi, Y., Chen, X., Elsasser, S., Stocks, B.B., Tian, G., Lee, B.H., Shi, Y., Zhang, N., de Poot, S.A., Tuebing, F., Sun, S., Vannoy, J., Tarasov, S.G., Engen, J.R., Finley, D., and Walters, K.J.** (2016). Rpn1 provides adjacent receptor sites for substrate binding and deubiquitination by the proteasome. *Science* **351**.
- Shibahara, T., Kawasaki, H., and Hirano, H.** (2004). Mass spectrometric analysis of expression of ATPase subunits encoded by duplicated genes in the 19S regulatory particle of rice 26S proteasome. *Arch. Biochem. Biophys.* **421**: 34–41.
- Sievers, F. and Higgins, D.G.** (2014). Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* **1079**: 105–16.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., and Higgins, D.G.** (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**: 539.
- Stadtmueller, B.M., Kish-Trier, E., Ferrell, K., Petersen, C.N., Robinson, H., Myszka, D.G., Eckert, D.M., Formosa, T., and Hill, C.P.** (2012). Structure of a proteasome Pba1-Pba2 complex: implications for proteasome assembly, activation, and biological function. *J. Biol. Chem.* **287**: 37371–82.

- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J.** (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*.
- Uechi, H., Hamazaki, J., and Murata, S.** (2014). Characterization of the testis-specific proteasome subunit alpha4s in mammals. *J. Biol. Chem.* **289**: 12365–74.
- Unverdorben, P., Beck, F., led, P., Schweitzer, A., Pfeifer, G., Plitzko, J.M., Baumeister, W., and Förster, F.** (2014). Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proc. Natl. Acad. Sci. U.S.A.* **111**: 5544–5549.
- Ustrell, V., Hoffman, L., Pratt, G., and Rechsteiner, M.** (2002). PA200, a nuclear proteasome activator involved in DNA repair. *EMBO J.* **21**: 3516–25.
- VanderLinden, R.T., Hemmis, C.W., Schmitt, B., Ndoja, A., Whitby, F.G., Robinson, H., Cohen, R.E., Yao, T., and Hill, C.P.** (2015). Structural basis for the activation and inhibition of the UCH37 deubiquitylase. *Mol. Cell* **57**: 901–11.
- Verma, R., Aravind, L., Oania, R., McDonald, W.H., Yates, J.R., Koonin, E.V., and Deshaies, R.J.** (2002). Role of the Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* **298**: 611–615.
- Vidal, M., Braun, P., Chen, E., Boeke, J.D., and Harlow, E.** (1996). Genetic characterization of a mammalian protein-protein interaction domain by using a yeast reverse two-hybrid system. *Proc. Natl. Acad. Sci. U.S.A.* **93**: 10321–6.
- Vierstra, R.D.** (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat. Rev. Mol. Cell Biol.* **10**: 385–97.
- Wang, D., Kalume, D., Pickart, C., Pandey, A., and Cotter, R.J.** (2006). Identification of protein ubiquitylation by electrospray ionization tandem mass spectrometric analysis of sulfonated tryptic peptides. *Anal. Chem.* **78**: 3681–7.
- Wani, P.S., Suppahia, A., Capalla, X., Ondracek, A., and Roelofs, J.** (2016). Phosphorylation of the C-terminal tail of proteasome subunit alpha7 is required for binding of the proteasome quality control factor Ecm29. *Sci. Rep.* **6**: 27873.
- Waskom, M.** (2016). Seaborn.
- Welk, V., Coux, O., Kleene, V., Abeza, C., Trumbach, D., Eickelberg, O., and Meiners, S.** (2016). Inhibition of Proteasome Activity Induces Formation of Alternative Proteasome Complexes. *J. Biol. Chem.* **291**: 13147–59.

- Wenger, C.D. and Coon, J.J.** (2013). A proteomics search algorithm specifically designed for high-resolution tandem mass spectra. *J. Proteome Res.* **12**: 1377–86.
- Yang, B.J., Han, X.X., Yin, L.L., Xing, M.Q., Xu, Z.H., and Xue, H.W.** (2016). Arabidopsis PROTEASOME REGULATOR1 is required for auxin-mediated suppression of proteasome activity and regulates auxin signalling. *Nat. Commun.* **7**: 11388.
- Yang, P., Fu, H., Walker, J., Papa, C.M., Smalle, J., Ju, Y.M., and Vierstra, R.D.** (2004). Purification of the Arabidopsis 26 S proteasome: biochemical and molecular analyses revealed the presence of multiple isoforms. *J. Biol. Chem.* **279**: 6401–13.
- Yao, T. and Cohen, R.E.** (2002). A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* **419**: 403–7.
- Yashiroda, H., Mizushima, T., Okamoto, K., Kameyama, T., Hayashi, H., Kishimoto, T., Niwa, S., Kasahara, M., Kurimoto, E., Sakata, E., Takagi, K., Suzuki, A., Hirano, Y., Murata, S., Kato, K., Yamane, T., and Tanaka, K.** (2008). Crystal structure of a chaperone complex that contributes to the assembly of yeast 20S proteasomes. *Nat. Struct. Mol. Biol.* **15**: 228–36.
- Yu, Y., Yu, Y., Smith, M., and Pieper, R.** (2014). A spinnable and automatable StageTip for high throughput peptide desalting and proteomics. *Protocol Exchange*.
- Zhang, X., Henriques, R., Lin, S.S., Niu, Q.W., and Chua, N.H.** (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* **1**: 641–6.

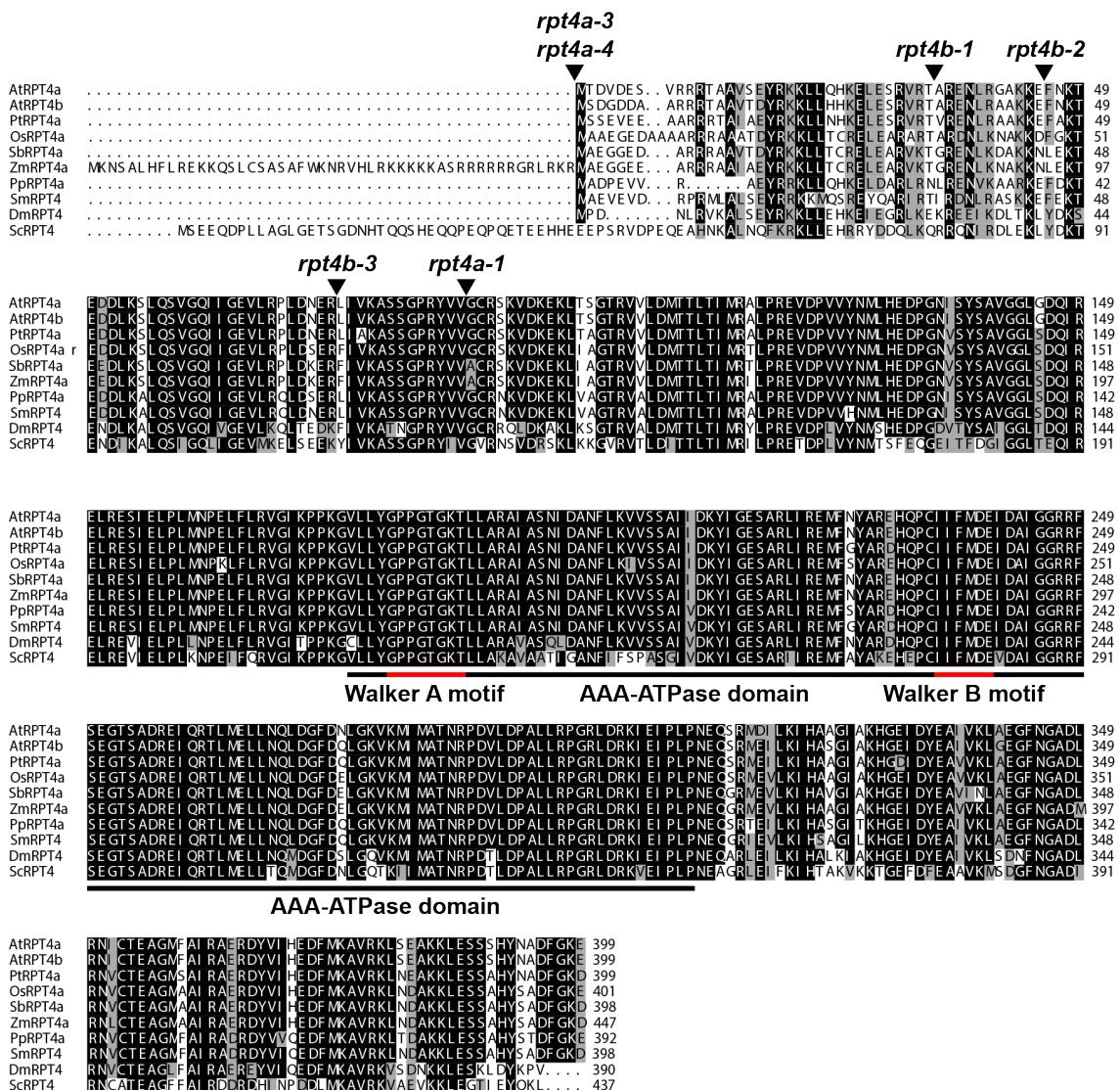


Figure 3.S1: Sequence Alignments of RPT4a and RPT4b. Sequence Alignments of *Arabidopsis* RPT4a RPT4b and related sequences was performed with ClustalW. Positions of T-DNA insertion mutants for both RPT4a and RPT4b were determined by sequencing T-DNA specific PCR products and are shown with black triangles. The AAA-ATPase domain is indicated with a black bar. The Walker A and B motifs contained in the AAA-ATPase domain are indicated in red. Species names are abbreviated as follows: At *Arabidopsis thaliana*, Pt *Populus trichocarpa*, Os *Oryza sativa*, Sb *Sorghum bicolor*, Zm- *Zea mays*, Pp *Physcomitrella patens*, Sm *Selaginella moellendorffii*, Dm *Drosophila melanogaster*, Sc- *Saccharomyces cerevisiae*.

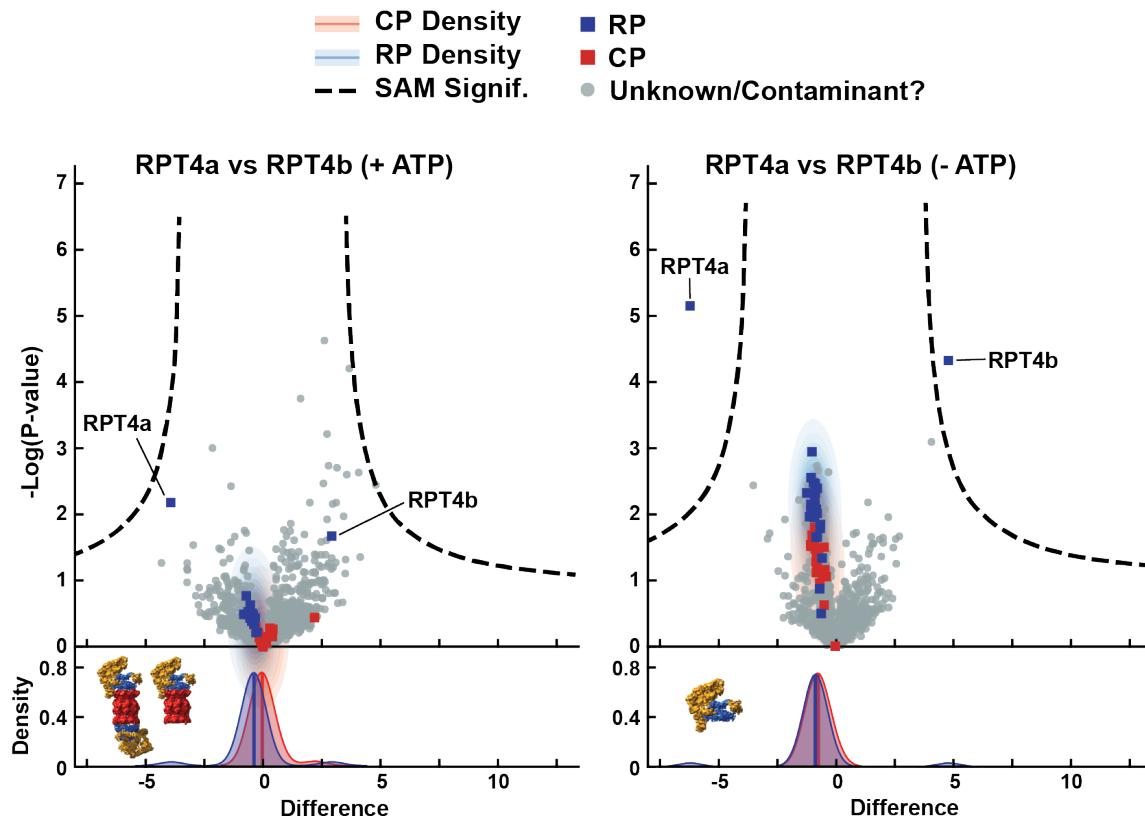


Figure 3.S2: Volcano Plots Comparing Protein Profiles of Proteasome Affinity Purifications Based on RPT4a or RPT4b subunits in the presence (+) or absence (-) of ATP. (A) In the presence of ATP, statistical analysis using significant analysis of microarray (SAM) test does not show enrichment for RPT4a in the RPT4a pulldown and RPT4b in the RPT4b pulldown. (B) In the absence of ATP, statistical analysis using a SAM test does show enrichment for RPT4a in the RPT4a pulldown and RPT4b in the RPT4b pulldown, but shows no other associations with central proteasome subunits or associated proteins. This suggests the complex assembles in a random fashion with respect to isoform incorporation.

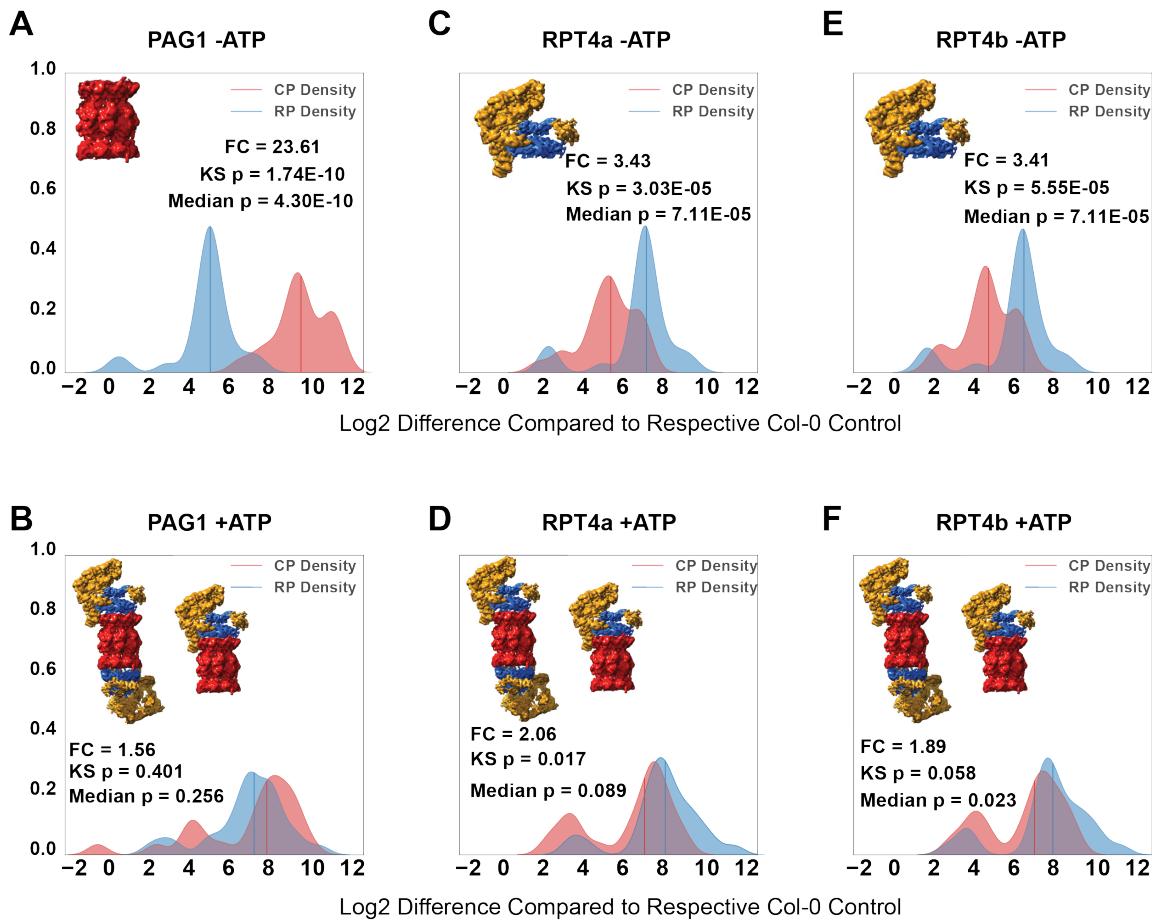


Figure 3.S3: Kernel Density Estimates (KDE) of the CP and RP distributions from label-free quantitative mass spectrometry show enrichment for the CP when purified via PAG1-FLAG, and for the RP when purified via FLAG-RPT4a/b. The X-axis shows difference in CP (red) or RP (blue) protein levels between affinity-purified samples and their respective mock affinity purification control, while the Y-axis is unit less with the area of each density corresponding to a total area of 1. Fold change (FC) CP over RP, or RP over CP was calculated based on the median of these distributions. Two *p*-values, one for a median test, and another for a Kolmogorov-Smirnov test, are shown, which test for differences in median and overall distribution respectively. Cryo-EM structures of the sub-complexes likely to be enriched in each affinity purification are shown, with free CP (A) and RP (C and E) in the absence of ATP, and in the presence of ATP half capped and doubly capped proteasomes (B, D, and F) adapted from (Beck *et al.*, 2012) **(A)** The CP is enriched over 20 fold when performing an affinity purification with ATP omitted from the purification buffer as compared to the RP. **(B)** The CP is only enriched 1.5 fold when affinity purified in the presence of 20 mM ATP. **(C)** The RP as compared to the CP is enriched approximately 3.5 fold in samples that were purified via FLAG-RPT4a in the absence of ATP. **(D)** In the presence of 20 mM ATP this RP enrichment decreases to only 2 fold for RPT4a. **(E)** In the absence of ATP the RP is enriched approximately 3.5 fold in samples affinity purified from RPT4b. **(F)** Enrichment for the RP decreases to about 2 fold in the presence of 20 mM ATP.

UMP1_a_A_t h - - - - - MES - - EKKI AHEI GGVKNDALRFGL - - H-
 UMP1_b_A_t h - - - - - MES - - QKKI AHEI GGMKNDALRFGL - - H-
 UMP1_c_A_t h - - - - -
 UMP1_Z_ma - - - - - MA - - SGS MKREI SET - HDTLRFGL - - NA
 UMP1_O_sa - - - - - MAS - - GGVVKKEI GGN - HDVLRFGV - - ND
 UMP1_M_tr - - - - - ME E - - SKTI PHQI GGVHTDVLRFGL - - P-
 UMP1_P_pa - - - - - MD - - GI QL PHDVLGS - HDSL LRHGL - - P-
 UMP1_S_ce - - - - - MNI VPQDTFKSQVSTDQ - - DKSVLSSAVPSL PDTL RQQEGGAV
 UMP1_K_la - - - - - MPMNI VPSPDFKS AVATDK - GCE HQS NAVASL PDVFR EQV - GAR
 MCS GETTI ASATPTQTAMNI VPPSNRRSALAA - - - GDATAASAAPALPDVLRERQR - AAR
 UMP1_D_ha - - - - - MSLKI VPDNNQHSSI NSTK - F - GE PAS HAP GL QDI LKTQE - GPL
 UMP1_D_me - - - - - MY - - QPSLK VQPAE - - VSVL NAT GRVGMP - - T - E - A-
 UMP1_X_la - - - - - MN - - TKAL GS QLKDSI PI GE LSTH YYGI QDTL RLTGF - T-
 UMP1_D_re - - - - - MN - - TRGL RS QLKDS VPVTGL CPQAGP YGVQDSL RRGF - S-
 UMP1_M_mu - - - - - MN - - ARGL GSEL KDSI PVAELS - ASGPFES HDL LRKGF - S-
 UMP1_H_sa - - - - - MN - - ARGL GSEL KDSI PVTELS - ASGPFES HDL LRKGF - S-
 consensus + + + + +

UMP1_a_A_t h GVKS NI GS HPLE SSYE SE KKSK EA LKRTVI AHAY GT AL PL KMDM DRQI LS RF - - -
 UMP1_b_A_t h GVKS DI LRS HPLE TAYE SGKQS C E MKR RV I THTYGA AL PL KMDL DRQI LS RF - - -
 UMP1_c_A_t h - - - - - ME PA PL KMDL DRP I L P RY GML DQYH
 UMP1_Z_ma GVKA DL APPHPL QSSI QE EAKF WADKKKF GTE A I YGS AL NI RKDL DAQI LS RF - - -
 UMP1_O_sa SVKGDL APPHPL QASVHKDAFKWADKKRFGAE A I YGS AF NI RKDL DAQI LS KF - - -
 UMP1_M_tr GVKS DI VGA HPLE LSSL QSVRGVE EAMRKQCKVNLYGA APL KE EL DRQI LS RF - - -
 UMP1_P_pa AFKS DVL PVHP VE VI QRSS KAMQA QT KRTL E RAY GS A P PMKMDI ERQI L TRF - - -
 UMP1_S_ce PLST QL NDRHPL EST LKNMETT QQR QME QYR QI FGG I AE PMKRTM E MEI VNRT - - -
 UMP1_K_la PLNT QL NDRHPL E S RVRN MDE T QH KRL E QYR QI FGG I AE PMKRV MEL KLV QNT - - -
 UMP1_E_go PLA AEL ADRHPL E S RL RDW DAC A QRT RLE QYR R I FGG I AE PVR RTM E L AL VS DA - - -
 UMP1_D_ha NLA SKI NNRHPL E GRI S NME E T QY QTR L E TY RRV F GAGE P I KRTM E L EL VE AT - - -
 UMP1_D_me NCL NQL AHV HRL RDSE L NYNE HQY NR NMQML RNHE GL GPV L KMGM E RFA AAR QV - - -
 UMP1_X_la SVQSE L MPSP HPL E L SE KNF QI NQEKVNFS T VRN I QGL HAP L KL HME F KAVKQV - - -
 UMP1_D_re SVK NELL PS HPL E L SE KNF QL NQDK MNF NT L RNI QGL HAP L KL QME YRAAK QI - - -
 UMP1_M_mu CVK NELL PS HPL E L SE KNF QL NQDK MNF ST L RNI QGL F APL KL QME F KAVQV - - -
 UMP1_H_sa CVK NELL PS HPL E L SE KNF QL NQDK MNF ST L RNI QGL F APL KL QME F KAVQV - - -
 consensus + + + + +

UMP1_a_A_t h - - - - - QRPPGP - I PSSM GL E VYT GAV DDF GFE DY L NDPR DS E T - FKP
 UMP1_b_A_t h - - - - - QRPPGP - I PSSM GL E VYT GAL DNF GFE DY L NDPR RE S E T - LKP
 UMP1_c_A_t h YE I PLNS DAYS A I RYE NRP TGP - I PSL MLD E VYT GAI DDF DFE DY L NVST P QI L - ENP
 UMP1_Z_ma - - - - - QRPPGA - L PSS L LG E ALT G S L DDF GFE DY L NMPQD S E G - FRQ
 UMP1_O_sa - - - - - QRPPGA - L PSS M L GYE ALT G S L DDF GFE DY L NL P QD S E S - FRA
 UMP1_M_tr - - - - - QRPPGV - I PSS M L G E L VT G T L DHF GFE DY L NDS R E S E T - FRP
 UMP1_P_pa - - - - - QRPT GL - I PSS M L G E S LT G E L D E F G F E D F L N DPK E S D T - FVP
 UMP1_S_ce - - - - - DF NPL - - STNGS I HRDI L L NKE CS I DWE DVYPGTGL QAS T MVG
 UMP1_K_la - - - - - DF NPL - - NO S N L H K D V L M N K E A S I DWE DVY PTS DF AS GMM V
 UMP1_E_go - - - - - EF APL - - GA A AL H L D V L N N T E A S L D W H D V Y P D E QQ - - -
 UMP1_D_ha - - - - - DF KP QLL GG S DS M HKD V L N K D A S V D W E D I Y R G G L E S G S - - HV
 UMP1_D_me - - - - - GRL P - F L S S S N F M D D V L T G R C D S I G F E D F M N L P E N S E H - MRQ
 UMP1_X_la - - - - - QRL P - F L P I S N I A L D T L R G T D E C I S F E D I L N D P S Q S E L - MGE
 UMP1_D_re - - - - - QRL P - F L P S S N L A L D T L R G S D D T I G F E D I L N D P V Q C E M - MGD
 UMP1_M_mu - - - - - HRL P - F L P S S N L S L D L R G N D E T I G F E D I L N D P S Q S E L - MGE
 UMP1_H_sa - - - - - QRL P - F L S S S N L S D L R G N D E T I G F E D I L N D P S Q S E V - MGE
 consensus + + + + +

UMP1_a_A_t h VDF - - - - - HHGME VRL GL S KGPI - - - - - YPS - - FM - - -
 UMP1_b_A_t h VDF - - - - - HHGME VRL GL S KGPA - - - - - SPS - - FM - - -
 UMP1_c_A_t h RCS NQL T S T T S WKL GL VSPKAQL NQVPCRS KQFI LYNAI KVL KL D S T C M F V Q P N A G E M I
 UMP1_Z_ma PDM - - - - - HHGME VRL GL S KGPI - - - - - CPS - - FS - - -
 UMP1_O_sa PDM - - - - - HHGME VRL GL S KGPV - - - - - CPS - - FN - - -
 UMP1_M_tr LDM - - - - - HHGME VRL GL S KGPV - - - - - YPS - - I I - - -
 UMP1_P_pa RDM - - - - - HES M VRL GL A KGPV - - - - - SRM - - FF - - -
 UMP1_S_ce DDV - - - - - HS K T E K Q L G I - - - - -
 UMP1_K_la DDV - - - - - HT K I E K R M G I - - - - -
 UMP1_E_go P D V - - - - - H A I E R S L R L - - - - -
 UMP1_D_ha QDF - - - - - H T E M E K K M G M - - - - -
 UMP1_D_me P - - - - - H A V V E K S L G I Y Q - - - - -
 UMP1_X_la P - - - - - H L M V E Y K L G L L - - - - -
 UMP1_D_re P - - - - - H I M T E Y K L G L L - - - - -
 UMP1_M_mu P - - - - - H V M V E H K L G L L - - - - -
 UMP1_H_sa P - - - - - H L M V E Y K L G L L - - - - -
 consensus + + + +

Figure 3.S4: Multiple Sequence Alignment for UMP1.

[Figure 3.S4 caption follows on next page]

UMP1_a_A. th	- - - - -
UMP1_b_A. th	- - - - -
UMP1_c_A. th	L AAP SSSG T KTL V E F N N K S A E T E P S D V K L G L D L V S S M K E E L D L G R N A P K Y F R V D F S P K
UMP1_Z. ma	- - - - -
UMP1_O. sa	- - - - -
UMP1_M. tr	- - - - -
UMP1_P. pa	- - - - -
UMP1_S. ce	- - - - -
UMP1_K. la	- - - - -
UMP1_E. go	- - - - -
UMP1_D. ha	- - - - -
UMP1_D. me	- - - - -
UMP1_X. la	- - - - -
UMP1_D. re	- - - - -
UMP1_M. mu	- - - - -
UMP1_H. sa	- - - - -
consensus
UMP1_a_A. th	- - - - -
UMP1_b_A. th	- - - - -
UMP1_c_A. th	D H T V C L E F G I L S R E F N R A H L R D G D K E K F T L K N V H K S K Q G E E A S L P T L I N L N K R D I N
UMP1_Z. ma	- - - - -
UMP1_O. sa	- - - - -
UMP1_M. tr	- - - - -
UMP1_P. pa	- - - - -
UMP1_S. ce	- - - - -
UMP1_K. la	- - - - -
UMP1_E. go	- - - - -
UMP1_D. ha	- - - - -
UMP1_D. me	- - - - -
UMP1_X. la	- - - - -
UMP1_D. re	- - - - -
UMP1_M. mu	- - - - -
UMP1_H. sa	- - - - -
consensus

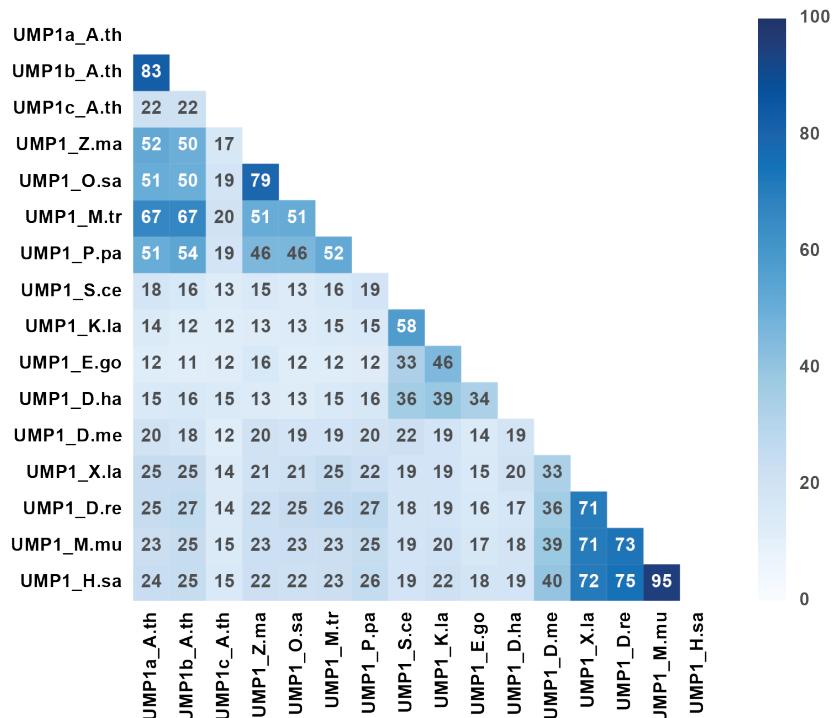


Figure 3.S4 (preceding page): Multiple Sequence Alignment for UMP1. A Clustal Omega (Sievers and Higgins, 2014) sequence alignment of UMP1 from plants, yeast, and animals visualized with the program BOXSHADE. A heat map showing each combination of protein percent identity with respective values shown in each cell obtained from the resultant Clustal Omega identity matrix. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaryomyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa, *Homo sapiens*.

Figure 3.S5: Multiple Sequence Alignment for PBAC1.

[Figure 3.S5 caption follows on next page]

PBAC1_A.th	T S A E K K A T T - H T G L V N L D Y F P S G S V I D G L I S A S L L S R C Q L K N I R G T L V I S W - P E - - - - F
PBAC1_Z.ma	T R A A R A R G G V G A A K A V Q A L V P P G S V V D G L I G A A V L A E G E I R G K A A S M V V T W - P A - - - - D
PBAC1_O.sa	T R A A R A R G R G G V G A A R G V A A L A P P G S V V D G L I G A A V L A E G E I M R G K A A S M V V T W - P E - - - - G
PBAC1_M.tr	S F A E R K K A N E E K L L G G L D Y Y P S G S V V D G L I G A A V L G R C Q I L K I R A S L C V T W - P Q - - - - F
PBAC1_P.pa	T S A S K K G K K - - N K D S I A P Y Y P S G S L V D G V A A A A L I T H Q C Q L R A L R A E L L M S W - P T - - - - A
PBAC1_S.ce	- - - - - S M K V M T E N E C T L Q P P E F I T G F I G S V L T Q L I V G P S K G L K F K C L V A P S E G P N G F
PBAC1_K.la	- - - - - E L K Q L S - - - Y L V P P E F I T G S I T N F I L A L - - - P - R K E L K V I V V P S E G P I G F
PBAC1_E.go	- - - - - G V K D L E - - - E L V P P E F I T G A L A A L V C A L - - - A P S G A G R V L V V Q S E G P H G F
PBAC1_D.ha	- - - - - Q T P D I Y A A I P S L I K P P H F I T G I G A S F N S Q I S L L R - - Q P R L M S L V L R S E G Q S G F
PBAC1_D.me	- - - - - A N Q K D I T E L E A P N F I A G V A A G V A S W R D Q M E L P V T S F V I Y - - T D - - - - K
PBAC1_X.la	T S E Y - - - - R E E V P C P L L E Q P N I V D G L I P A A V L S H C Q V L G I P A V F Y R C Y - - T D - - - - I
PBAC1_D.re	T S T Y - - - - T H T V T C P P L E Q P N I C S G L A A A V L S H C Q V H Q I S A V L Y Q C Y - - S D - - - - V
PBAC1_M.mu	T Q T F - - - - K D A L C C P L L E Q P N I V H D L I S A A V L S Y C O V V K I P A V L Y L C Y - - T D - - - - V
PBAC1_H.sa	T Q N F - - - - K D S A C C P L L E Q P N I V H D L I P A A V L S Y C O V V K I P A I L Y L C Y - - T D - - - - V
consensus	+ + . + . + + . + + + + + + . + + . + . + . + + . + +
PBAC1_A.th	G P S V I R F A G A L L K K I V P S I D V A S V N K D L E K N - - S S R T G L K K - D T W L - - D S D L Y T
PBAC1_Z.ma	S R P A E F - - - G V M R R V A A E L G V D T G K - - - - - A - A A R V S - G R P D - - L D S L Y T
PBAC1_O.sa	A R P A E F - - - G V M R R V A A E L G V D P A K - - - - - A A A G R V A - G R A E - - L D A L Y T
PBAC1_M.tr	D S D V V L L L K D L L R G L G E F - E F G F S D - - D E - - V F K F G R S K - D H V F - - Q S H L Y I
PBAC1_P.pa	D T S V V N Q L A S A L Q N V C Q Q A - - L A N N V I I D F - - I A A A K S I A - G R R L L S D S E I Y M
PBAC1_S.ce	E K I L S L S D M G S L V D L C G Q W I G F - - E - P S R Y S E E C - Y R L W R C D S A A I G A Q S G L Y I
PBAC1_K.la	E K Y N F P T V D A L V D E M S K L L C S Y P S Q - - T E Y Y V S E C - L K L W K L D E C T - T T Q G G L Y I
PBAC1_E.go	E K H N L T V V D A L V E R V R Q I L D - - - Q - - P P E Y V A A C - T R H W K L Q A G S - Q F Q G G L Y L
PBAC1_D.ha	E K I I D T D A F V D A S F V L N E L L V G T S D Q - - G - N Y L K Q V S S T V R K I N G - - Y S N S I G M Y I
PBAC1_D.me	L P L D - - - A T A A Q P V I K V L K A A G V A - C S S Y I P P R K E S - - - - - S Y L Y M
PBAC1_X.la	S K L D - - - S V T I K A F R P L L S S G S L S - R L A A D S A N I Q E T L R K T - V K L N E I Q S N L Y I
PBAC1_D.re	L H P D - - - S A S M Q M F A A T L S S V - - - L K L E Q S L S A D V L Q R V - T R V S E I Q S N L Y T
PBAC1_M.mu	M K L D - - - R V T V E A F K P L L S S R S L K - C L V K N I P E S T E I L K K L - M T T N E I Q S N I Y T
PBAC1_H.sa	M K L D - - - L I T V E A F K P I L S T R S L K - G L V K N I P Q S T E I L K K L - M T T N E I Q S N I Y T
consensus	. + + + . + . + . + + + + + + . + * +

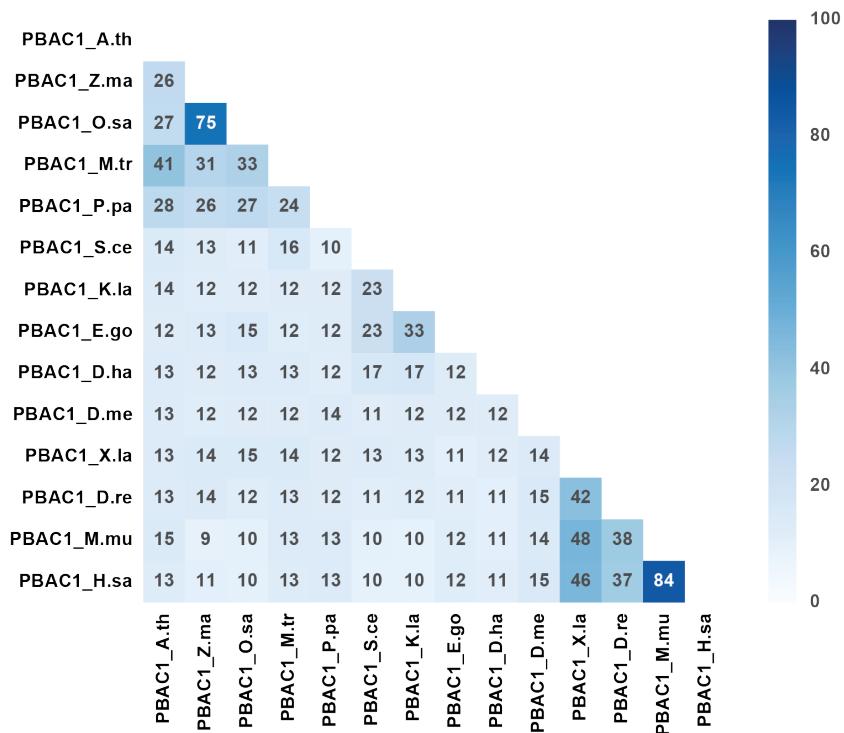


Figure 3.S5 (preceding page): Multiple Sequence Alignment for PBAC1. A Clustal Omega (Sievers and Higgins, 2014) sequence alignment of PBAC1 from plants, yeast, and animals visualized with the program BOXSHADE. A heat map showing each combination of protein percent identity with respective values shown in each cell obtained from the resultant Clustal Omega identity matrix. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaromyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa, *Homo sapiens*.

Figure 3.S6: Multiple Sequence Alignment for PBAC2.

[Figure 3.S6 caption follows on next page]

PBAC2_A. t h	FSAF	- - - - -	K A R G L K M T C L C Y C S E G D N I P E A F L L A E A - - S S K N G L T P D K F H - G E E G
PBAC2_Z. m z	F S A C	- - - - -	K A K G L K V S C V L C Y C S E G D N M P E S F Q L A E A - - I C K L L G H G P E K F H - G N G S
PBAC2_O. s a	F S A C	- - - - -	K A K G L K V T C V L C Y C S E G D N M P E S F Q L A E A - - A C K L V A Q G P E Q F H - G N G S
PBAC2_M. t r	F S F L	- - - - -	K A K G L K V T C L L C Y C S E G D N M P E S F Q L A D A - - A C K L L R L S H P N S - G I E G
PBAC2_P. pa	L A C C	- - - - -	K A Q G L K V V C I L C F A E G D N V P D A F L I A D G - L H R L L C H E K K E P Q - G S G
PBAC2_S. c e	F Q I	L N A S - Q R P K A L R S I K Y C S C L A N E G D N S L D S Q Q F L Q W I I S Q K V I K N A P P - - - - - I	
PBAC2_K. l a	H H L L Y Q	- - - - -	I T Y L S K V V H E G D N S W D A F Q I C E K - - L A E L Q D L I - - - - - Q
PBAC2_E. go	Y H L L H S P L A R S A D F A E L S Y V S I F I H E G D N N E D A L L C T E - - L P T F V E V P R - - - - - L		
PBAC2_D. ha	T Y L H T G D L S	- - - - -	S S - - D E F V N V L V S Y V E G D N F Y D G E H L A N K T I E A - L G V S P - - - - - V
PBAC2_D. me	F K A F	- - - - -	- - D E Q K V P V M L F R Y L L E G D N S T D A S L V R E - - L N E L C E D F L Q L R N G G D G S
PBAC2_X. l a	Y N D C	- - - - -	C L E D L Q M A V V L K F C S E G D N V P D A F S L L N Q - - V N E W W H L V E P T N - - - - - G D
PBAC2_D. re	Y T D C	- - - - -	C T E D V P M A V V L I F C S E G D N I P D A F A L I N C - - L I N D W L H L L E K - - - P T Q G S
PBAC2_M. mu	Y D E S	- - - - -	C S K E I Q M A V V L K F V S E G D N I P D A V S L V E Y - - L N E W L Q I K P C N D G P M A S A
PBAC2_H. s a	Y D E S	- - - - -	C S K E I Q M A V V L K F V S E G D N I P D A L G L V E Y - - L N E W L Q I K P L S D D P T V S A
consensus	+	+ . + . + + + * * * + + + + . + . + . + + . +

PBAC2_A. t h	GKWWI	PYSWKS	MYGAPPDMS	MF-	- - -
PBAC2_Z. m rz	NGWI	VPLSWKS	VYGPPPDMS	I F-	- - -
PBAC2_O. s a	NGMTI	PLSWKS	VYGPPPDLS	I F-	- - -
PBAC2_M. t r	GKWR	IPLSWMSVY	GPPPDVS	I F-	- - -
PBAC2_P. pa	VTWKI	PLSWTT	VYGPPPD	D T MFC	- - -
PBAC2_S. c e	VKFVRP	I S WQGAY	GMA DARDK	FVDL	- YN
PBAC2_K. l a	GRITAPV	S WKGVY	GARPI	AS GFDQGI	Y S
PBAC2_E. go	ETLRTPI	S WKGVYGS	RRI	PDGYDEGLYI	Y
PBAC2_D. ha	KNM	I RPRSWL	G VYGD RQVPNAME	E EGLFG	-
PBAC2_D. me	FKLT	T VPKS	WNLLF	GNDVT ELL	F - - -
PBAC2_X. l a	VKWKSPRS	W RL L	F GS GL PPAI	F - - -	-
PBAC2_D. re	VQMRVPS	PSWKL	L F GS GI PPL	L F - - -	-
PBAC2_M. mu	L PWKI	PSSWRL	L F GS GL PPA	L F - - -	-
PBAC2_H. s a	S RWKI	PSSWRL	L F GS GL PPA	L F - - -	-
consensus	.	++++*	+**	+. +*++.

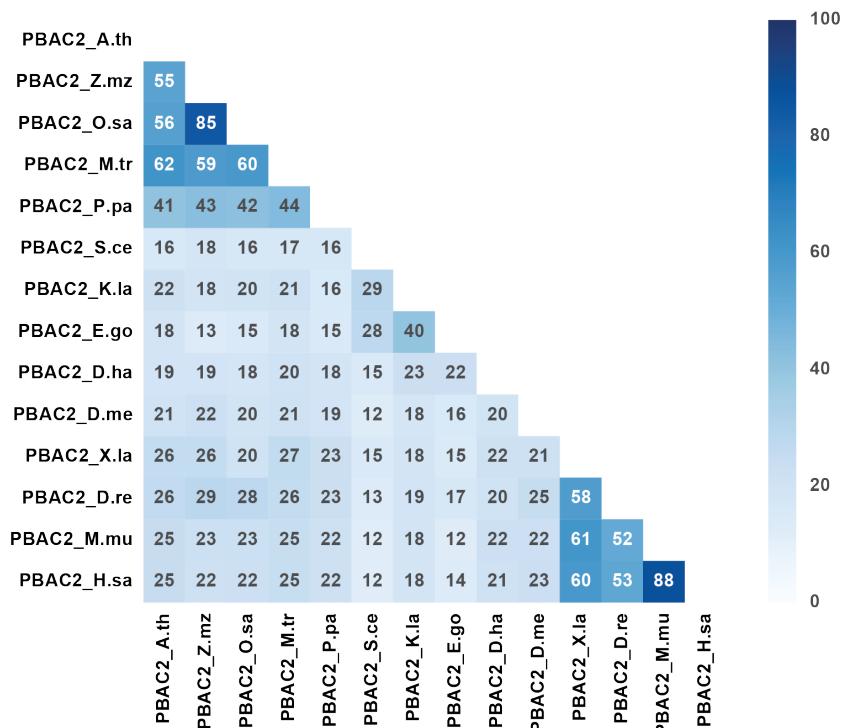


Figure 3.S6 (preceding page): Multiple Sequence Alignment for PBAC2. A Clustal Omega (Sievers and Higgins, 2014) sequence alignment of PBAC2 from plants, yeast, and animals visualized with the program BOXSHADE. A heat map showing each combination of protein percent identity with respective values shown in each cell obtained from the resultant Clustal Omega identity matrix. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaromyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa, *Homo sapiens*.

Figure 3.S7: Multiple Sequence Alignment for PBAC3.

[Figure 3.S7 caption follows on next page]

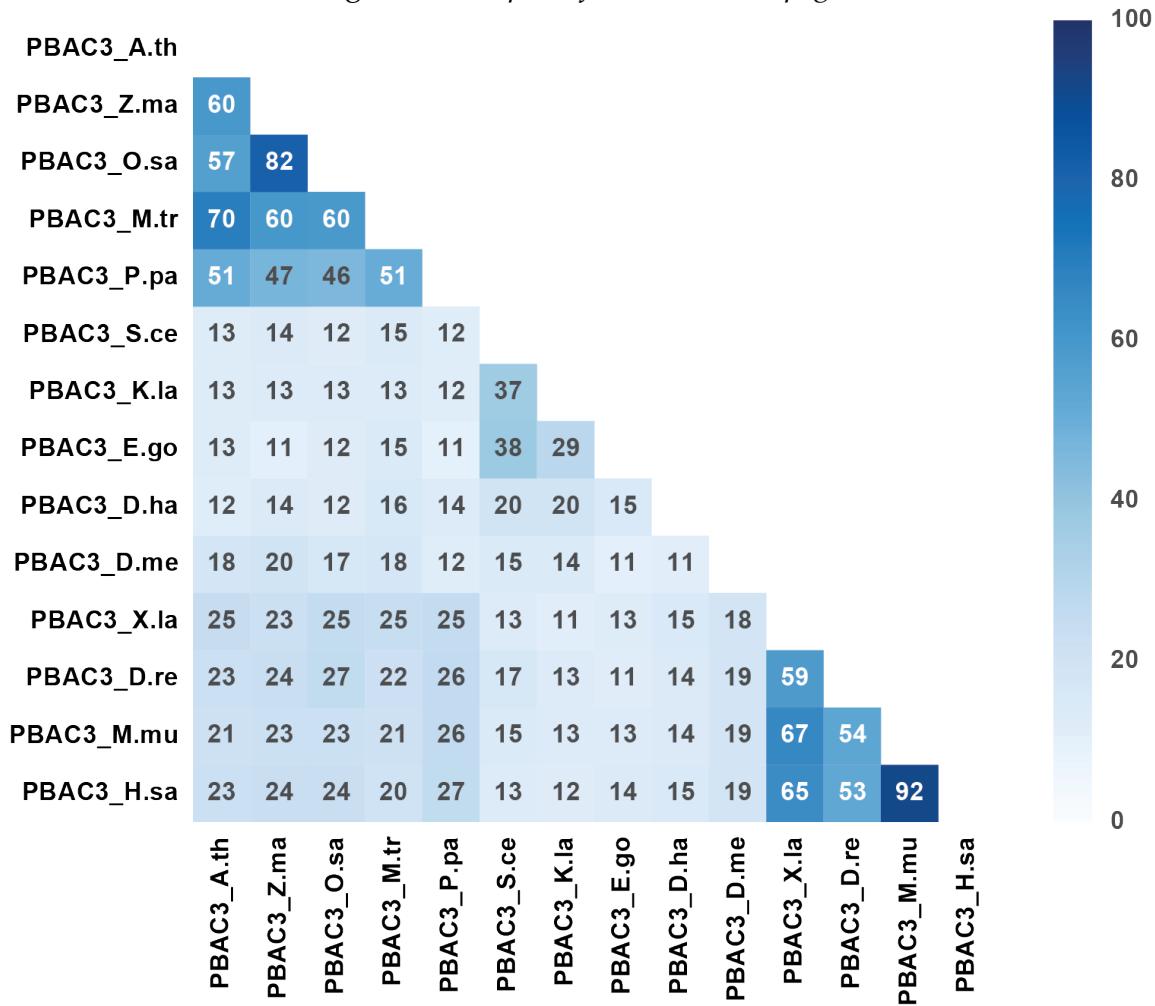


Figure 3.S7 (preceding page): Multiple Sequence Alignment for PBAC3. A Clustal Omega (Sievers and Higgins, 2014) sequence alignment of PBAC3 from plants, yeast, and animals visualized with the program BOXSHADE. A heat map showing each combination of protein percent identity with respective values shown in each cell obtained from the resultant Clustal Omega identity matrix. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaromyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa, *Homo sapiens*.

Figure 3.S8: Multiple Sequence Alignment for PBAC4.

[Figure 3.S8 caption follows on next page]

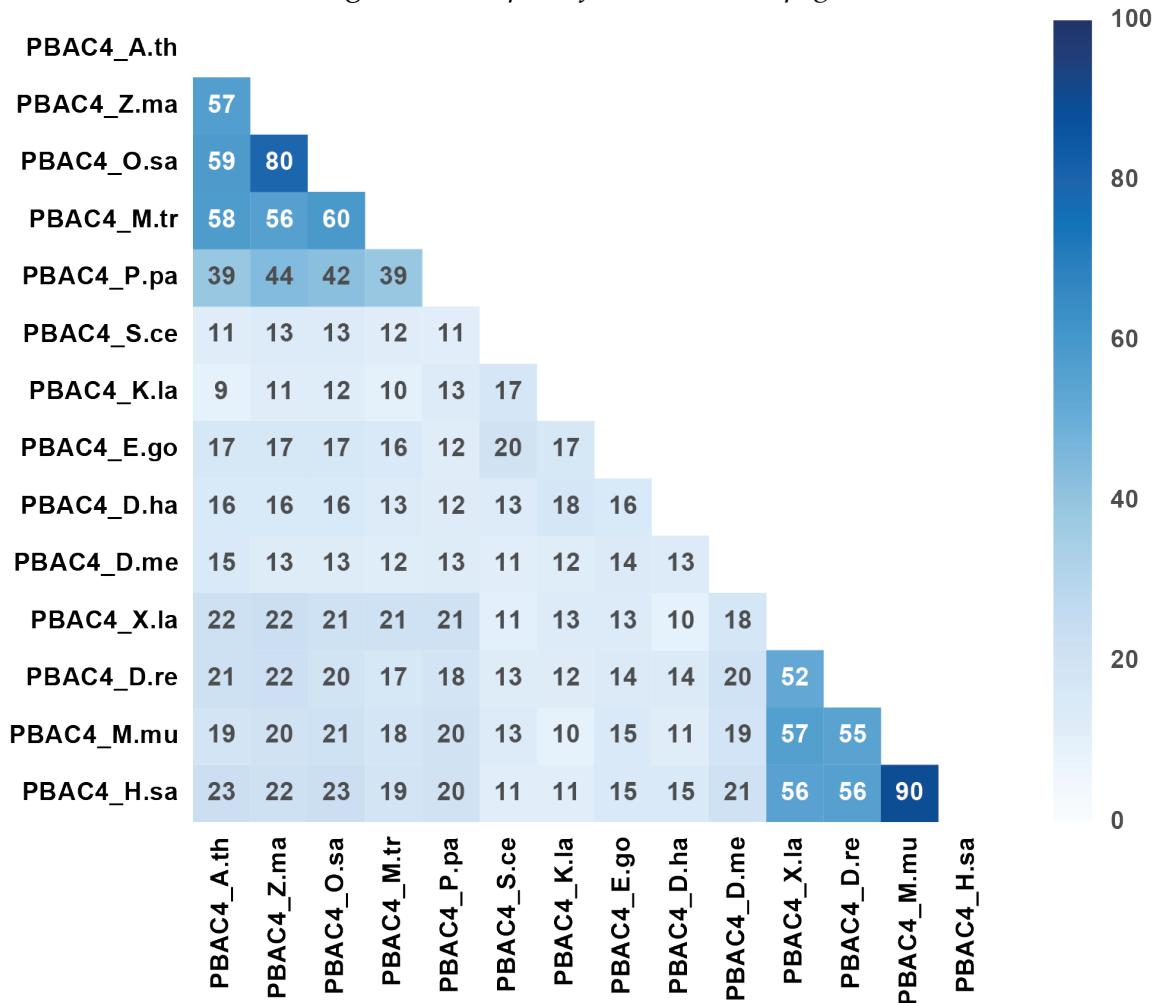


Figure 3.S8 (preceding page): Multiple Sequence Alignment for PBAC4. A Clustal Omega (Sievers and Higgins, 2014) sequence alignment of PBAC4 from plants, yeast, and animals visualized with the program BOXSHADE. A heat map showing each combination of protein percent identity with respective values shown in each cell obtained from the resultant Clustal Omega identity matrix. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaromyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa, *Homo sapiens*.

[Figure 3.S9 caption follows on next page]

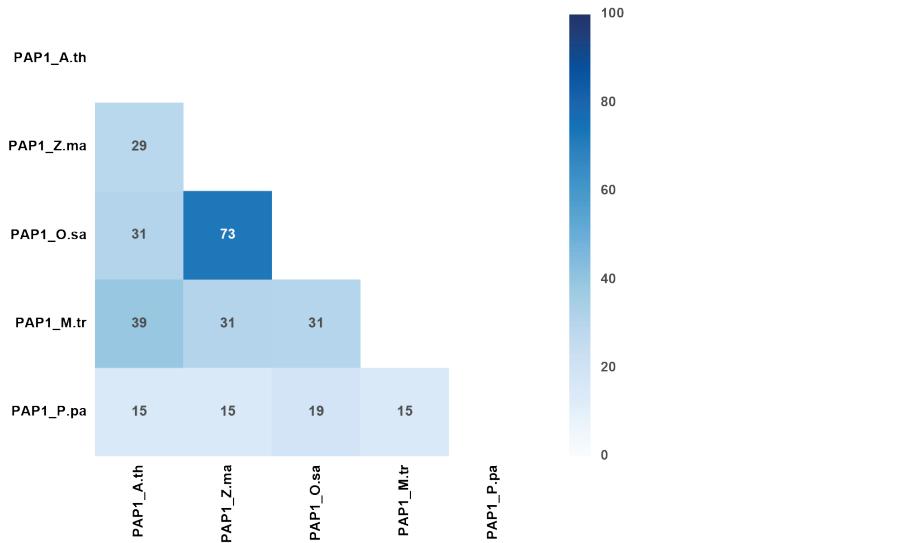
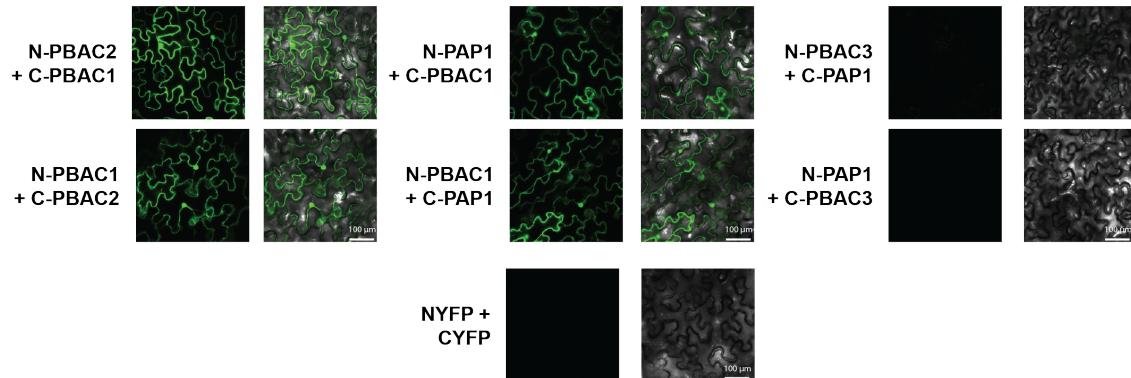


Figure 3.S9 (preceding page): Multiple Sequence Alignment for PAP1. A Clustal Omega (Sievers and Higgins, 2014) sequence alignment of PAP1 from plants, yeast, and animals visualized with the program BOXSHADE. A heat map showing each combination of protein percent identity with respective values shown in each cell obtained from the resultant Clustal Omega identity matrix. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaryomyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa *Homo sapiens*.

[Figure 3.S10 caption follows on next page]

A



B

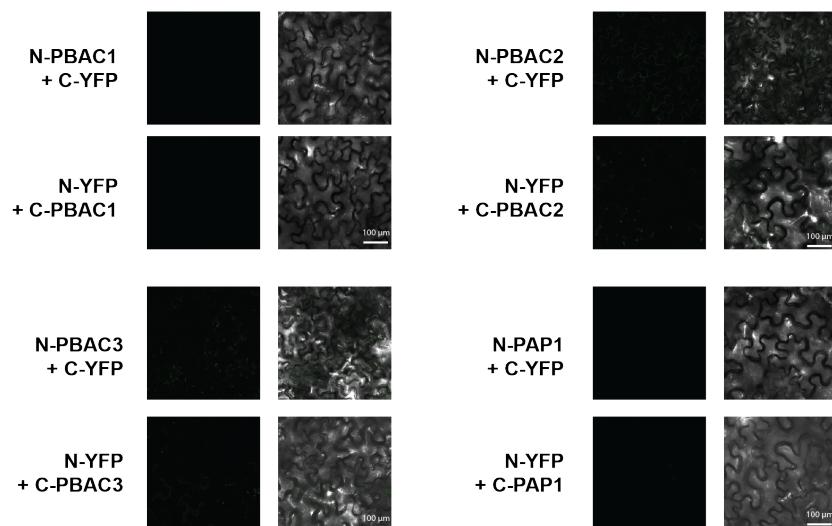


Figure 3.S10 (preceding page): Full BiFC Panel and Controls. BiFC analysis showing both GFP channel and a merged bright-field and GFP channel. **(A)** A repeat of Figure 6D is shown for reference **(B)** Additional controls combining N-YFP and C-YFP with their respective N and C terminal YFP-fused PBAC1, PBAC2, PBAC3, and PAP1.

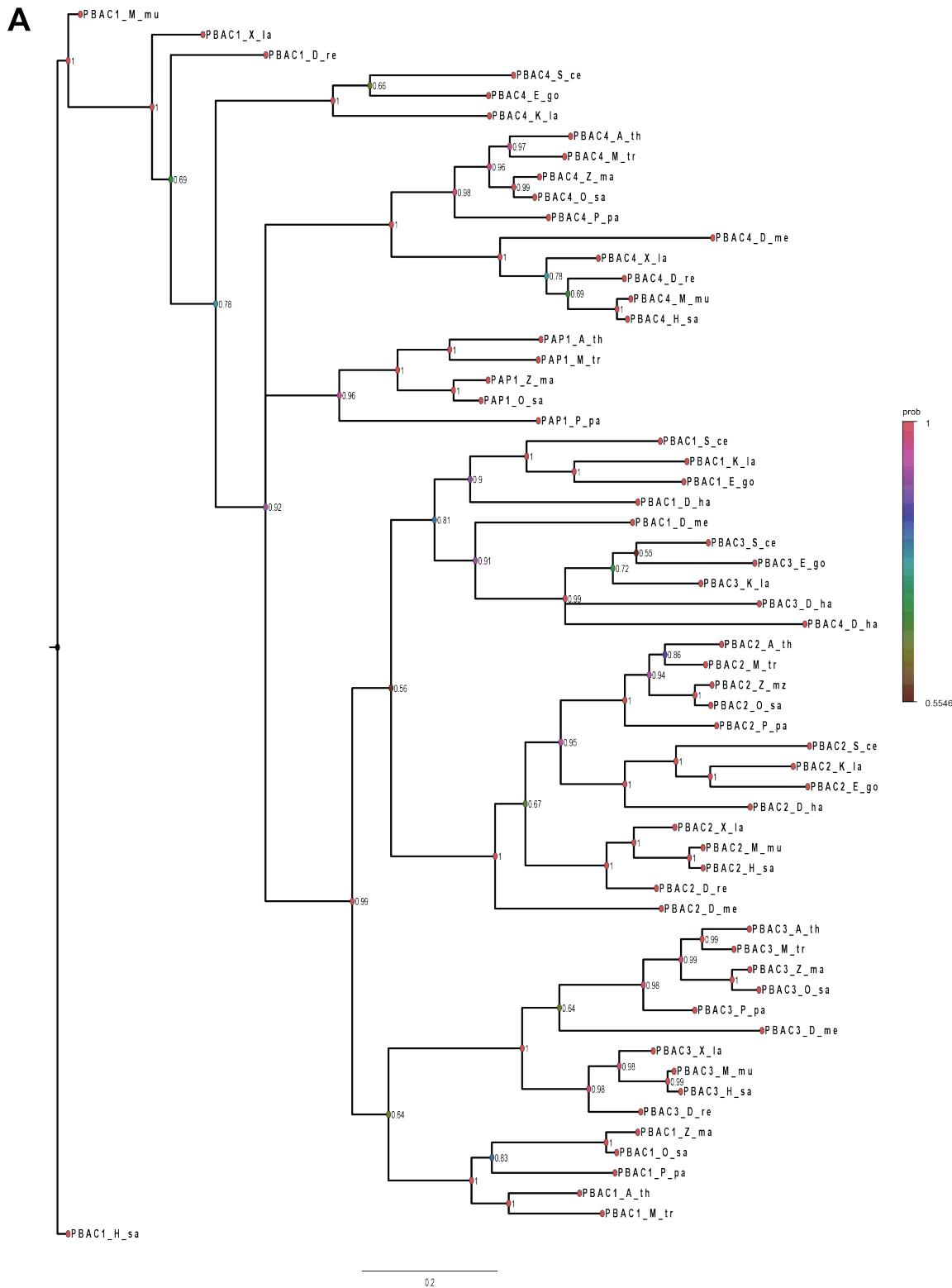


Figure 3.S11: Phylogenetic Analysis of PAP1 and PBAC1-4.

[Figure 3.S11 caption follows on next page]

B

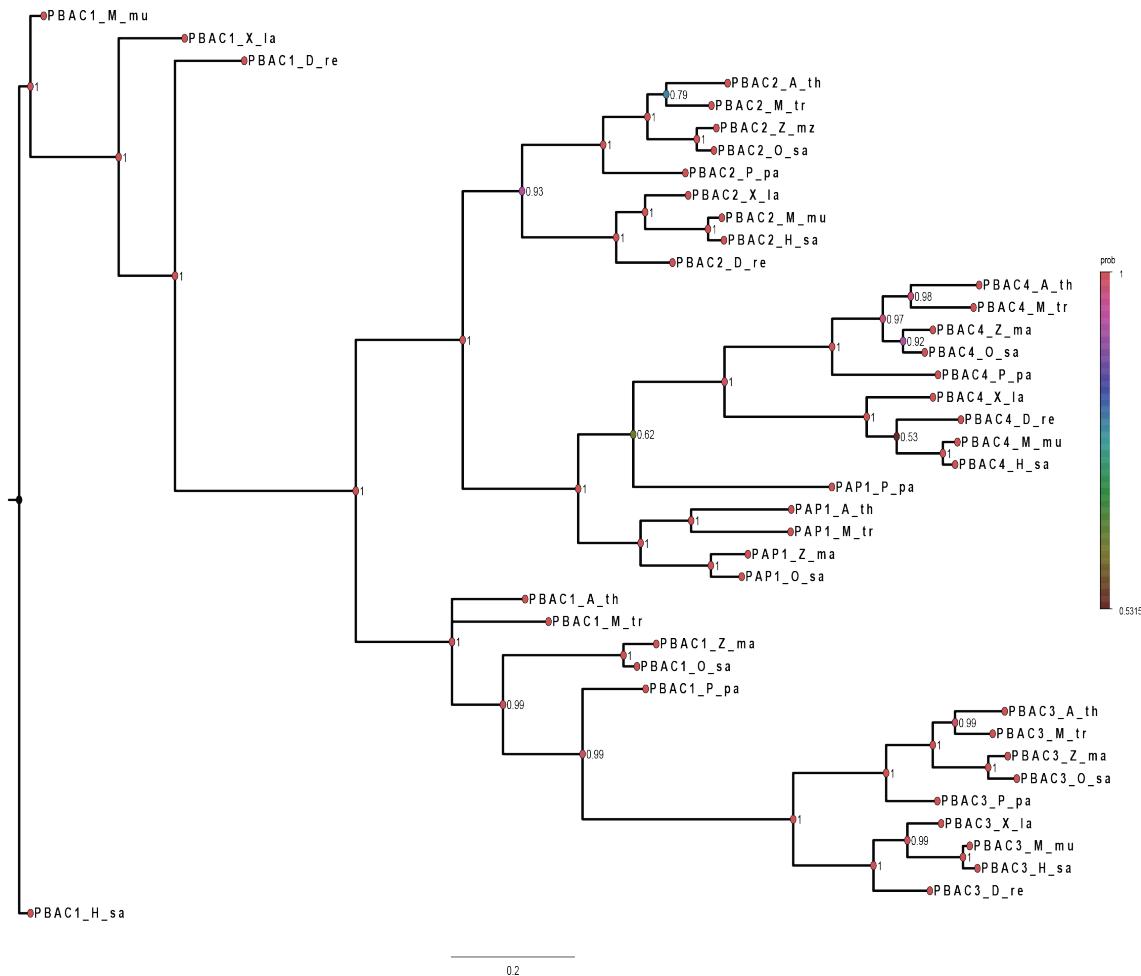


Figure 3.S11 (preceding page): Phylogenetic Analysis of PAP1 and PBAC1-4.. Bayesian phylogenetic analysis of PAP1 and PBAC1-4 from sequence alignments shown in Figure S5-S9. Trees were generated with MrBayes (Ronquist *et al.*, 2012) and then visualized with FigTree. Nodes are color-coded with clade probabilities (probability of observing this grouping in the set of sampled trees), and a color key is shown on the right. Species names are abbreviated as follows: A.th *Arabidopsis thaliana*, Z.ma *Zea mays*, Os *Oryza sativa*, M.tr *Medicago truncatula*, P.pa *Physcomitrella patens*, S.ec- *Saccharomyces cerevisiae* K.la *Kluyveromyces lactis*, E.go *Eremothecium gossypii*, D.ha *Debaryomyces hansenii*, D.me *Drosophila melanogaster*, X.la *Xenopus laevis*, D.re *Danio rerio*, M.mu *Mus musculus*, H.sa, *Homo sapiens*. **(A)** Phylogeny generated with all sequences from yeast, animals, and plants. Clear grouping is observed between the yeast, animal, and putative plant PBAC2 orthologs into a single clade. Grouping was observed for PBAC3 between the animal, and putative plant sequences; however, yeast PBAC3 was not found in this clade. Clear grouping is observed for PBAC4 between the animal, and putative plant sequences; however, yeast PBAC4 was not found in this clade. Animal, yeast, and plant PBAC1 did not clearly group, likely due to considerable sequence divergence. **(B)** Phylogeny generated when omitting the yeast sequences and only utilizing animal and plant sequences. Clear grouping is observed between the putative plant orthologs PBAC2-4 with their respective animal orthologs into separate clades. Animal PBAC1 does not group with the putative plant specific PBAC1 orthologs, likely due to considerable sequence divergence from animal PBAC1.

Table 3.S1: Oligonucleotide Primer Sequences Used in this Study. A table of primers used throughout this chapter is provided for reference.

Table 3.S2: Contaminants Removed from MaxQuants ProteinGroups.txt file.
 A list of common contaminants that were removed from MaxQuants ProteinGroups.txt file for the proteomics experiments involving PAG1-FLAG, and FLAG-RPT4a/b affinity purified proteasomes. ATG Identifiers, Symbols and descriptions from The *Arabidopsis* Information Resource v 10 (TAIR10) are provided.

ATG Identifier	Symbols	TAIR10 Description
AT1G03880.1	CRU2, CRB	cruciferin 2
AT3G44300.1	NIT2, AtNIT2	nitrilase 2
AT3G44310.3	NIT1, ATNIT1, NIT1	nitrilase 1
AT3G44310.2	NIT1, ATNIT1, NIT1	nitrilase 1
AT4G27140.1	SESA1, AT2S1	seed storage albumin 1
AT4G27160.1	AT2S3, SESA3	seed storage albumin 3
AT4G27170.1	SESA4, AT2S4	seed storage albumin 4
AT4G28520.1	CRU3, CRC	cruciferin 3
AT5G44120.3	CRA1, ATCRA1, CRU1	RmlC-like cupins superfamily protein
AT5G54740.1	SESA5	seed storage albumin 5

Table 3.S3: InterProScan analysis of *Arabidopsis* protein sequences for PBAC1-4. InterProScan Signatures from both PFAM and Panther databases with their respective expectation-values (E-values) are shown for reference (Jones *et al.*, 2014). While PBAC1 (previously identified as a likely ortholog for PAC1 (Kus-mierczyk *et al.*, 2011)) contains a Proteasome Assembly Chaperone (PAC) 2 domain shown with a star, it is of considerably lower expectation-value (E-value) than PBAC2 which contains a Proteasome Assembly Chaperone (PAC) 2 domain. PBAC3 contains a PAC3 domain, and PBAC4 contains a PAC4 domain. PAP1 contains no domains of known function but does contain a conserved domain annotated in the Panther database (Mi *et al.*, 2005).

Protein	AGI Identifier	InterproScan Signature	Annotation	E-value	Start Residue	End Residue
PBAC1	AT3G25545	PF09754 PTHR37227	Proteasome Assembly Chaperone 2* Family Not Named	3.4E-06	60	247
PBAC2	AT3G18940	PF09754 PTHR12970	Proteasome Assembly Chaperone 2 Member 5-induced protein 1 class3	1.6E-116	2	267
PBAC3	AT5G14710	PF10178 PTHR31051	Proteasome Assembly Chaperone 3 Proteasome Assembly Chaperone 3	1.0E-29	18	249
PBAC4	AT1G48170	PF16093 PTHR33559	Proteasome Assembly Chaperone 4 Family Not Named	4.8E-102	3	172
PAP1	AT3G07640	PTHR37221	Family Not Named	1.6E-26	37	120
				2.9E-45	1	124
				1.4E-17	47	120
				9.5E-56	14	153
				8.1E-109	2	241

Table 3.S4: Iterative PSI-BLAST Analysis using Human PAC1-4 to Identify PBAC1-4. Iterative PSI-BLAST Analysis was performed on NCBI's non-redundant (nr) database using the human PAC1-4 sequences using the default parameters. PBAC1-4 were the top hits in *Arabidopsis* for their respective PAC1-4 search sequences. ATG identifiers, RefSeq identifiers, Number of PSI-BLAST iterations, and expectation values (E-value) are listed for each analysis. PBAC1 required a second iteration to be identified, in which all sequences above the default expectation threshold were included in the second search.

Human	Arabidopsis	ATG Identifier	Matching Accession	Iteration	E-value
PAC1	PBAC1	AT3G25545	NP_566770.1	2	2.0E-04
PAC2	PBAC2	AT3G18940	NP_566621.1	1	1.0E-21
PAC3	PBAC3	AT5G14710	NP_196975.1	1	5.0E-08
PAC4	PBAC4	AT1G48170	NP_564522.1	1	4.0E-06

Table S4 Proteasome associated proteins identified after PAG1-based affinity purification under both plus and minus ATP Conditions. A table of proteins identified as significantly enriched at FDR 0.01 in the indicated affinity preparations. Shown are the proteins ATG identifiers, descriptions from *The Arabidopsis Information Resource version 10* (TAIR10), symbols from TAIR10, average MS/MS count, average sequence coverage, fold change compared to control using the MaxLFQ value from MaxQuant, and *p*-values obtained from the volcano plots in Figure 4A and B.

Table S5 Proteasome associated proteins identified after RPT4a-based affinity purification under both plus and minus ATP conditions. A table of proteins identified as significantly enriched at FDR 0.01 in the indicated affinity preparations. Shown are the proteins ATG identifiers, descriptions from *The Arabidopsis Information Resource version 10* (TAIR10), symbols from TAIR10, average MS/MS count, average sequence coverage, fold change compared to control using the MaxLFQ value from MaxQuant, and *p*-values obtained from the volcano plots in Figure 4C and D.

Table S6 Proteasome associated proteins identified after RPT4b-based affinity purification under both plus and minus ATP conditions. A table of proteins identified as significantly enriched at FDR 0.01 in the indicated affinity preparations.

[Table 3.S5 caption follows on next page] []

PAG1-FLAG -ATP

ATG	TAIR10 Description	Symbol(s)	Proteasome Annotation	Average		Fold Change Compared to Control	P-Value
				Average MS/MS Coverage Count*	Average Sequence (%)		
AT3G18940.1	clast3-related		PBAC2	41.5	56.2	2734.5	1.21E-06
AT3G07640.1	unknown protein		PAP1	49.8	59.9	2462.4	6.05E-06
AT3G25545.1	unknown protein		PBAC1	58.2	68.4	2319.3	2.73E-07
AT3G13330.1	proteasome activating protein 200	PA200	PA200	54.3	30.1	846.7	3.02E-05
AT5G38650.1	Proteasome maturation factor UMP1		Ump1b	11.5	67.4	279.2	1.26E-05
AT1G67250.1	Proteasome maturation factor UMP1		Ump1a	8.7	56.0	277.0	1.96E-05
AT3G53970.1	proteasome inhibitor-related		PTRE1	6.3	29.4	240.6	9.99E-06
AT1G23520.1	Domain of unknown function (DUF220)			8.7	3.0	165.3	1.02E-03
AT1G01320.2	Tetratricopeptide repeat (TPR)-like superfamily protein			2.0	1.7	156.0	4.31E-07
AT1G48170.1	unknown protein		PBAC4	19.8	79.5	138.4	5.30E-02
AT5G14710.1	CONTAINS InterPro DOMAIN/s: Proteasome assembly chaperone 3 (InterPro:IPR018788)		PBAC3	7.8	39.4	126.3	5.79E-05
AT3G13090.1	multidrug resistance-associated protein 8	ATMRP8, ABCC6, MRP8		0.0	0.5	86.8	1.01E-01
AT2G47360.1	unknown protein			3.2	2.6	64.5	3.39E-03
AT2G26780.1	ARM repeat superfamily protein		ECM29	3.5	9.7	35.6	1.91E-02
AT4G17100.1	CONTAINS InterPro DOMAIN/s: Endoribonuclease XendoU (InterPro:IPR018998)			0.2	4.4	26.0	6.55E-05
AT5G27960.1	AGAMOUS-like 90			0.2	6.3	25.1	2.71E-02
AT1G21530.1	AMP-dependent synthetase and ligase family protein			0.2	3.1	24.4	1.22E-05
AT1G23780.1	F-box family protein			2.0	8.6	21.5	1.63E-04
AT3G26020.2	Protein phosphatase 2A regulatory B subunit family protein			0.8	2.7	14.5	2.53E-04
AT2G17200.1	ubiquitin family protein	DSK2		2.5	9.4	13.7	1.10E-02
ATCG00680.1	photosystem II reaction center protein B	PSBB		0.0	2.2	13.4	3.07E-04
AT5G55220.1	trigger factor type chaperone family protein			0.2	2.3	13.2	5.56E-03
AT5G13450.1	delta subunit of Mt ATP synthase	ATP5		0.0	4.5	8.5	9.55E-04
AT1G79530.1	glyceraldehyde-3-phosphate dehydrogenase of plastid 1	GAPCP-1		3.3	17.6	7.4	3.56E-04

PAG1-FLAG +ATP

ATG	TAIR10 Description	Symbol(s)	Proteasome Annotation	Average		Fold Change Compared to Control	P-Value
				Average MS/MS Coverage Count*	Average Sequence (%)		
AT3G25545.1	unknown protein		PBAC1	33.7	40.7	2130.4	8.05E-06
AT3G18940.1	clast3-related		PBAC2	12.2	39.6	1340.9	1.45E-05
AT2G26780.1	ARM repeat superfamily protein		ECM29	0.0	0.8	908.0	5.36E-04
AT3G13330.1	proteasome activating protein 200	PA200	PA200	10.8	11.8	755.5	1.20E-05
AT3G07640.1	unknown protein		PAP1	21.0	27.4	446.8	1.09E-03
AT5G38650.1	Proteasome maturation factor UMP1		Ump1b	6.2	45.4	326.9	1.47E-04
AT1G48170.1	unknown protein		PBAC4	8.3	41.1	291.7	6.90E-05
AT1G67250.1	Proteasome maturation factor UMP1		Ump1a	4.7	36.7	247.3	2.59E-04
AT2G47360.1	unknown protein			2.7	2.6	197.7	1.08E-05
AT3G13090.1	multidrug resistance-associated protein 8	ATMRP8, ABCC6, MRP8		0.0	1.3	184.1	1.15E-02
AT5G14710.1	CONTAINS InterPro DOMAIN/s: Proteasome assembly chaperone 3 (InterPro:IPR018788)		PBAC3	2.8	24.5	162.2	4.32E-05
AT3G48670.2	XH/XS domain-containing protein	IDN2, RDM12		1.5	2.2	113.5	3.33E-04
AT2G28070.1	ABC-2 type transporter family protein			0.8	1.8	105.9	2.27E-04
AT2G14170.2	aldehyde dehydrogenase 6B2	ALDH6B2		0.8	1.0	100.8	5.43E-05
AT5G27960.1	AGAMOUS-like 90	AGL90		0.0	6.3	80.4	4.59E-04
AT1G74260.1	purine biosynthesis 4	PUR4		0.3	0.9	62.9	1.70E-02
AT3G53970.1	proteasome inhibitor-related	PTRE1		2.3	14.4	62.7	3.29E-03
AT5G66420.1	CONTAINS InterPro DOMAIN/s: Uncharacterised conserved protein UCP033271 (InterPro:IPR008322), TIM-barrel signal transduction protein, predicted (InterPro:IPR009215)			0.2	1.5	57.2	1.03E-03
AT3G26020.2	Protein phosphatase 2A regulatory B subunit family protein			0.5	2.3	54.7	1.55E-04
AT2G37650.1	GRAS family transcription factor			0.3	0.7	46.7	2.71E-04

Shown are the proteins ATG identifiers, descriptions from *The Arabidopsis Information Resource version 10 (TAIR10)*, symbols from TAIR10, average MS/MS count, average sequence coverage, fold change compared to control using the MaxLFQ value from MaxQuant, and *p*-values obtained from the volcano plots in Figure 4E and F.

Table S7 dNSAF values determined by Morpheus Spectral Counter used to generate the heatmap in Figure 6C. Proteins as annotated in the heatmap are listed with each slice and treatment status as treated or untreated with MG132. dNSAF values as determined by Morpheus Spectral Counter (Gemperline *et al.*, 2016) are listed for each slice.

Table S8 Fold Change dNSAF values comparing Slice 1 and Smear from Figure 6. Fold change values were calculated for both treated and untreated samples by dividing the dNSAF values for Smear / Slice 1. #DIV/0! Is listed as a divide by zero error, in that those proteins were detected in the smear but not in Slice 1. Larger values were color coded by red in Microsoft Excel, and smaller values were color coded by blue.

Table 3.S5 (preceding page): PAG1 Associated

COLOPHON

This document was typeset with L^AT_EX . It is based on the University of Wisconsin dissertation template created by William C. Benton (available at <https://github.com/willb/wi-thesis-template>).